Discrete time approximation of fully nonlinear HJB equations via BSDEs with nonpositive jumps

Idris Kharroubi, Nicolas Langrené, Huyên Pham

Discrete time approximation of fully nonlinear HJB equations via BSDEs with nonpositive jumps

Idris Kharroubi*
CEREMADE, CNRS UMR 7534,
Université Paris Dauphine
and CREST,
kharroubi at ceremade.dauphine.fr

Nicolas Langrené
Laboratoire de Probabilités et Modèles Aléatoires,
Université Paris Diderot
and EDF R&D
langrene at math.univ-paris-diderot.fr

Huyên Pham
Laboratoire de Probabilités et Modèles Aléatoires,
Université Paris Diderot
and CREST-ENSAE
pham at math.univ-paris-diderot.fr

November 22, 2013

Abstract

We propose a new probabilistic numerical scheme for fully nonlinear equation of Hamilton-Jacobi-Bellman (HJB) type associated to stochastic control problem, which is based on the Feynman-Kac representation in [13] by means of control randomization and backward stochastic differential equation with nonpositive jumps. We study a discrete time approximation for the minimal solution to this class of BSDE when the time step goes to zero, which provides both an approximation for the value function and for an optimal control in feedback form. We obtained a convergence rate without any ellipticity condition on the controlled diffusion coefficient. Explicit implementable scheme based on Monte-Carlo simulations and empirical regressions, associated error analysis, and numerical experiments are performed in the companion paper [14].

Key words: Discrete time approximation, Hamilton-Jacobi-Bellman equation, nonlinear degenerate PDE, optimal control, backward stochastic differential equations.

MSC Classification: 65C99, 60J75, 49L25.

*The research of the author benefited from the support of the French ANR research grant LIQUIRISK (ANR-11-JS01-0007).

1 Introduction

Let us consider the fully nonlinear generalized Hamilton-Jacobi-Bellman (HJB) equation:

$$\begin{cases} \dfrac{\partial v}{\partial t} + \sup\limits_{a\in A}\big[b(x,a).D_x v + \dfrac{1}{2}\mathrm{tr}(\sigma\sigma^{\intercal}(x,a)D_x^2 v) + f(x,a,v,\sigma^{\intercal}(x,a)D_x v)\big] & = & 0, \quad \text{on } [0,T)\times\mathbb{R}^d, \\ v(T,x) & = & g, \quad \text{on } \mathbb{R}^d. \end{cases} \tag{1.1}$$

In the particular case where $f(x,a)$ does not depend on v and $D_x v$, this partial differential equation (PDE) is the dynamic programming equation for the stochastic control problem:

$$v(t,x) \quad = \quad \sup_{\alpha}\mathbb{E}\Big[\int_t^T f(X_s^{\alpha},\alpha_s)ds + g(X_T^{\alpha})\Big|X_t^{\alpha} = x\Big], \tag{1.2}$$

with controlled diffusion in $\mathbb{R}^d$:

$$dX_t^{\alpha} \quad = \quad b(X_t^{\alpha},\alpha_t)dt + \sigma(X_t^{\alpha},\alpha_t)dW_t,$$

and where α is an adapted control process valued in a compact space A of $\mathbb{R}^q$. Numerical methods for parabolic partial differential equations (PDEs) are largely developed in the literature, but remain a big challenge for fully nonlinear PDEs, like the HJB equation (1.1), especially in high dimensional cases. We refer to the recent paper [10] for a review of some deterministic and probabilistic approaches.

In this paper, we propose a new probabilistic numerical scheme for HJB equation, relying on the following Feynman-Kac formula for HJB equation obtained by randomization of the control process α. We consider the minimal solution (Y,Z,U,K) to the backward stochastic differential equation (BSDE) with nonpositive jumps:

$$\begin{cases} Y_t & = & g(X_T) + \int_t^T f(X_s,I_s,Y_s,Z_s)ds + K_T - K_t \\ & & \quad - \int_t^T Z_s dW_s - \int_t^T\int_A U_s(a)\tilde{\mu}(ds,da), \qquad 0\le t\le T, \\ U_t(a) & \le & 0, \end{cases} \tag{1.3}$$

with a forward Markov regime-switching diffusion process (X,I) valued in $\mathbb{R}^d\times A$ given by:

$$\begin{aligned} X_t &= X_0 + \int_0^t b(X_s,I_s)ds + \int_0^t \sigma(X_s,I_s)dW_s \\ I_t &= I_0 + \int_{(0,t]}\int_A (a - I_{s^-})\mu(ds,da). \end{aligned}$$

Here W is a standard Brownian motion, $\mu(dt,da)$ is a Poisson random measure on $[0,\infty)\times A$ with finite intensity measure $\lambda(da)$ of full topological support on A, and compensated measure $\tilde{\mu}(dt,da) = \mu(dt,da) - \lambda(da)dt$. Assumptions on the coefficients b,σ,f,g will be detailed in the next section, but we emphasize the important point that no degeneracy condition on the controlled diffusion coefficient σ is imposed. It is proved in [13] that the minimal solution to this class of BSDE is related to the HJB equation (1.1) through the relation $Y_t = v(t,X_t)$.

The purpose of this paper is to provide and analyze a discrete-time approximation scheme for the minimal solution to (1.3), and thus an approximation scheme for the HJB equation. In the non-constrained jump case, approximations schemes for BSDE have been studied in the papers [12], [7], which extended works in [8], [21] for BSDEs in a Brownian framework. The issue is now to deal with the nonpositive jump constraint in (1.3), and we propose a discrete time approximation scheme of the form:

$$\left\{\begin{array}{rcl} \bar Y^\pi_T \;=\; \bar{\mathcal{Y}}^\pi_T & = & g(\bar X^\pi_T) \\ \bar{\mathcal{Z}}^\pi_{t_k} & = & \mathbb{E}\Big[\bar Y^\pi_{t_{k+1}} \frac{W_{t_{k+1}}-W_{t_k}}{t_{k+1}-t_k}\big|\mathcal{F}_{t_k}\Big] \\ \bar{\mathcal{Y}}^\pi_{t_k} & = & \mathbb{E}\Big[\bar Y^\pi_{t_{k+1}}\big|\mathcal{F}_{t_k}\Big] + (t_{k+1}-t_k)\, f(\bar X^\pi_{t_k}, I_{t_k}, \bar{\mathcal{Y}}^\pi_{t_k}, \bar{\mathcal{Z}}^\pi_{t_k}) \\ \bar Y^\pi_{t_k} & = & \underset{a\in A}{\operatorname{ess\,sup}}\, \mathbb{E}\Big[\bar{\mathcal{Y}}^\pi_{t_k}\big|\mathcal{F}_{t_k}, I_{t_k}=a\Big], \quad k=0,\ldots,n-1, \end{array}\right. \tag{1.4}$$

where $\pi = \{t_0 = 0 < \ldots < t_k < \ldots < t_n = T\}$ is a partition of the time interval $[0,T]$, with modulus $|\pi|$, and $\bar X^\pi$ is the Euler scheme of X (notice that I is perfectly simulatable once we know how to simulate the distribution $\lambda(da)/\int_A \lambda(da)$ of the jump marks). The interpretation of this scheme is the following. The first three lines in (1.4) correspond to the standard scheme $(\bar{\mathcal{Y}}^\pi, \bar{\mathcal{Z}}^\pi)$ for a discretization of a BSDE with jumps (see [7]), where we omit here the computation of the jump component. The last line in (1.4) for computing the approximation $\bar Y^\pi$ of the minimal solution Y corresponds precisely to the minimality condition for the nonpositive jump constraint and should be understood as follows. By the Markov property of the forward process (X,I), the solution $(\mathcal{Y},\mathcal{Z},\mathcal{U})$ to the BSDE with jumps (without constraint) is in the form $\mathcal{Y}_t = \vartheta(t,X_t,I_t)$ for some deterministic function ϑ. Assuming that ϑ is a continuous function, the jump component of the BSDE, which is induced by a jump of the forward component I, is equal to $\mathcal{U}_t(a) = \vartheta(t,X_t,a) - \vartheta(t,X_t,I_{t^-})$. Therefore, the nonpositive jump constraint means that: $\vartheta(t,X_t,I_{t^-}) \geq \underset{a\in A}{\operatorname{ess\,sup}}\,\vartheta(t,X_t,a)$. The minimality condition is thus written as:

$$Y_t \;=\; v(t,X_t) \;=\; \underset{a\in A}{\operatorname{ess\,sup}}\,\vartheta(t,X_t,a) \;\;=\;\; \underset{a\in A}{\operatorname{ess\,sup}}\,\mathbb{E}[\mathcal{Y}_t|X_t, I_t = a],$$

whose discrete time version is the last line in scheme (1.4).

In this work, we mainly consider the case where $f(x,a,y)$ does not depend on z, and our aim is to analyze the discrete time approximation error on Y, where we split the error between the positive and negative parts:

$$\mathrm{Err}^\pi_+(Y) \;:=\; \Big(\max_{k\leq n-1} \mathbb{E}\Big[\big(Y_{t_k} - \bar Y^\pi_{t_k}\big)_+^2\Big]\Big)^{\frac12}, \qquad \mathrm{Err}^\pi_-(Y) \;:=\; \Big(\max_{k\leq n-1} \mathbb{E}\Big[\big(Y_{t_k} - \bar Y^\pi_{t_k}\big)_-^2\Big]\Big)^{\frac12}.$$

We do not study directly the error on Z, and instead focus on the approximation of an optimal control for the HJB equation, which is more relevant in practice. It appears that the maximization step in the scheme (1.4) provides a control in feedback form $\{\hat a(t_k, \bar X^\pi_{t_k}), k \leq n-1\}$, which approximates the optimal control with an estimated error bound. The analysis of the error on Y proceeds as follows. We first introduce the solution $(Y^\pi, \mathcal{Y}^\pi, \mathcal{Z}^\pi, \mathcal{U}^\pi)$ of a discretely jump-constrained BSDE. This corresponds formally to BSDEs for which the nonpositive jump constraint operates only a finite set of times, and should be viewed as the analog of discretely reflected BSDEs defined in [1] and [6] in the context of the approximation for reflected BSDEs. By combining BSDE methods and PDE approach with comparison principles, and further with the shaking coefficients method of Krylov [17] and Barles, Jacobsen [4], we prove the monotone convergence of this discretely jump-constrained BSDE towards the minimal solution to the BSDE with nonpositive jump constraint, and obtained a convergence rate without any ellipticity condition on the diffusion coefficient σ. We next focus on the approximation error between the discrete time scheme in (1.4) and the discretely jump-constrained BSDE. The standard argument for studying rate of convergence of such error consists in getting an estimate of the error at time t_k: $\mathbb{E}[|Y^\pi_{t_k} - \bar{Y}^\pi_{t_k}|^2]$ in function of the same estimate at time t_{k+1}, and then conclude by induction together with classical estimates for the forward Euler scheme. However, due to the supremum in the conditional expectation in the scheme (1.4) for passing from $\bar{\mathcal{Y}}^\pi$ to $\bar{Y}^\pi$, which is a nonlinear operation violating the law of iterated conditional expectations, such argument does not work anymore. Instead, we consider the auxiliary error control at time t_k:

$$\mathcal{E}^\pi_k(\mathcal{Y}) \quad := \quad \mathbb{E}\Big[\operatorname*{ess\,sup}_{a\in A} \mathbb{E}_{t_1,a}\big[\dots \operatorname*{ess\,sup}_{a\in A} \mathbb{E}_{t_k,a}\big[|\mathcal{Y}^\pi_{t_k} - \bar{\mathcal{Y}}^\pi_{t_k}|^2\big]\dots\big]\Big],$$

where $\mathbb{E}_{t_k,a}[.]$ denotes the conditional expectation $\mathbb{E}[.|\mathcal{F}_{t_k}, I_{t_k} = a]$, and we are able to express $\mathcal{E}^\pi_k(\mathcal{Y})$ in function of $\mathcal{E}^\pi_{k+1}(\mathcal{Y})$. We define similarly an error control $\mathcal{E}^\pi_k(X)$ for the forward Euler scheme, and prove that it converges to zero with a rate $|\pi|$. Proceeding by induction, we then obtain a rate of convergence $|\pi|$ for $\mathcal{E}^\pi_k(\mathcal{Y})$, and consequently for $\mathbb{E}[|Y^\pi_{t_k} - \bar{Y}^\pi_{t_k}|^2]$. This leads finally to a rate $|\pi|^{\frac{1}{2}}$ for $\mathrm{Err}^\pi_-(Y)$, $|\pi|^{\frac{1}{10}}$ for $\mathrm{Err}^\pi_+(Y)$, and so $|\pi|^{\frac{1}{10}}$ for the global error $\mathrm{Err}^\pi(Y) = \mathrm{Err}^\pi_+(Y) + \mathrm{Err}^\pi_-(Y)$. Moreover, in the case where $f(x,a)$ does not depend on y (i.e. the case of standard HJB equation and stochastic control problem), we obtain a better rate of order $|\pi|^{\frac{1}{6}}$ by relying on a stochastic control representation of the discretely jump-constrained BSDE, and by using a convergence rate result in [16] for the approximation of controlled diffusion by means of piece-wise constant policies. Anyway, our result improves the convergence rate of the mixed Monte-Carlo finite difference scheme proposed in [10], where the authors obtained a rate $|\pi|^{\frac{1}{4}}$ on one side and $|\pi|^{\frac{1}{10}}$ on the other side under a nondegeneracy condition.

We conclude this introduction by pointing out that the above discrete time scheme is not yet directly implemented in practice, and requires the estimation and computation of the conditional expectations together with the supremum. Actually, simulation-regression methods on basis functions defined on $\mathbb{R}^d \times A$ appear to be very efficient, and provide approximate optimal controls in feedback forms via the maximization operation in the last step of the scheme (1.4). We postpone this analysis and illustrations with several numerical tests arising in superreplication of options under uncertain volatility and correlation in a companion paper [14]. Notice that since it relies on the simulation of the forward process (X, I), our scheme does not suffer the curse of dimensionality encountered in finite difference scheme or controlled Markov chains methods (see [18], [5]), and takes advantage of the high-dimensional properties of Monte-Carlo methods.

The remainder of the paper is organized as follows. In Section 2, we state some useful auxiliary error estimate for the Euler scheme of the regime switching forward process. We introduce in Section 3 discretely jump-constrained BSDE and relate it to a system of integro-partial differential equations. Section 4 is devoted to the convergence of discretely jump-constrained BSDE to the minimal solution of BSDE with nonpositive jumps. We provide in Section 5 the approximation error for our discrete time scheme, and as a byproduct an estimate for the approximate optimal control in the case of classical HJB equation associated to stochastic control problem.

2 The forward regime switching process

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space supporting d-dimensional Brownian motion W, and a Poisson random measure $\mu(dt, da)$ with intensity measure $\lambda(da)dt$ on $[0, \infty) \times A$, where A is a compact set of $\mathbb{R}^q$, endowed with its Borel tribe $\mathcal{B}(A)$, and λ is a finite measure on $(A, \mathcal{B}(A))$ with full topological support. We denote by $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ the completion of the natural filtration generated by (W, μ), and by $\mathcal{P}$ the σ-algebra of $\mathbb{F}$-predictable subsets of $\Omega \times \mathbb{R}_+$.

We fix a finite time horizon $T > 0$, and consider the solution (X, I) on $[0, T]$ of the regime-switching diffusion model:

$$\left\{ \begin{array}{rcl} X_t & = & X_0 + \int_0^t b(X_s, I_s)ds + \int_0^t \sigma(X_s, I_s)dW_s \\ I_t & = & I_0 + \int_{(0,t]} \int_A (a - I_{s^-})\mu(ds, da), \end{array} \right. \tag{2.1}$$

where $(X_0, I_0) \in \mathbb{R}^d \times A$, $b : \mathbb{R}^d \times A \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \times A \to \mathbb{R}^{d \times d}$, are measurable functions, satisfying the Lipschitz condition:

(H1) There exists a constant L_1 such that

$$|b(x, a) - b(x', a')| + |\sigma(x, a) - \sigma(x', a')| \quad \leq \quad L_1\big(|x - x'| + |a - a'|\big),$$

for all $x, x' \in \mathbb{R}^d$ and $a, a' \in A$. The assumption **(H1)** stands in force throughout the paper, and in this section, we shall denote by C_1 a generic positive constant which depends only on L_1, T, (X_0, I_0) and $\lambda(A) < \infty$, and may vary from lines to lines. Under **(H1)**, we have the existence and uniqueness of a solution to (2.1), and in the sequel, we shall denote by $(X^{t,x,a}, I^{t,a})$ the solution to (2.1) starting from (x, a) at time t.

Remark 2.1 We do not make any ellipticity assumption on σ. In particular, some lines and columns of σ may be equal to zero, and so there is no loss of generality by considering that the dimension d of X and W are equal. □

We first study the discrete-time approximation of the forward process. Denoting by $(T_n, \iota_n)_n$ the jump times and marks associated to μ, we observe that I is explicitly written as:

$$I_t \quad = \quad I_0 \mathbb{1}_{[0,T_1)}(t) + \sum_{n \geq 1} \iota_n \mathbb{1}_{[T_n, T_{n+1})}(t), \quad 0 \leq t \leq T,$$

where the jump times $(T_n)_n$ evolve according to a Poisson distribution of parameter $\lambda := \int_A \lambda(da) < \infty$, and the i.i.d. marks $(\iota_n)_n$ follow a probability distribution $\bar{\lambda}(da) := \lambda(da)/\lambda$. Assuming that one can simulate the probability distribution $\bar{\lambda}$, we then see that the pure jump process I is perfectly simulated. Given a partition $\pi = \{t_0 = 0 < \ldots < t_k < \ldots t_n = T\}$ of $[0,T]$, we shall use the natural Euler scheme $\bar X^\pi$ for X, defined by:

$$
\begin{aligned}
\bar X^\pi_0 &= X_0 \\
\bar X^\pi_{t_{k+1}} &= \bar X^\pi_{t_k} + b(\bar X^\pi_{t_k}, I_{t_k})(t_{k+1}-t_k) + \sigma(\bar X^\pi_{t_k}, I_{t_k})(W_{t_{k+1}} - W_{t_k}),
\end{aligned}
$$

for $k = 0, \ldots, n-1$. We denote as usual by $|\pi| = \max_{k \leq n-1}(t_{k+1}-t_k)$ the modulus of π, and assume that $n|\pi|$ is bounded by a constant independent of n, which holds for instance when the grid is regular, i.e. $(t_{k+1}-t_k) = |\pi|$ for all $k \leq n-1$. We also define the continuous-time version of $\bar X^\pi$ by setting:

$$
\bar X^\pi_t = \bar X^\pi_{t_k} + b(\bar X^\pi_{t_k}, I_{t_k})(t-t_k) + \sigma(\bar X^\pi_{t_k}, I_{t_k})(W_t - W_{t_k}), \quad t \in [t_k, t_{k+1}], \ k < n.
$$

By standard arguments, see e.g. [15], one can obtain under **(H1)** the L^2-error estimate for the above Euler scheme:

$$
\mathbb{E}\Big[\sup_{t\in[t_k,t_{k+1}]} \big|X_t - \bar X^\pi_{t_k}\big|^2\Big] \leq C_1|\pi|, \quad k < n.
$$

For our purpose, we shall need a stronger result, and introduce the following error control for the Euler scheme:

$$
\mathcal{E}^\pi_k(X) := \mathbb{E}\Big[\operatorname*{ess\,sup}_{a\in A} \mathbb{E}_{t_1,a}\big[\ldots \operatorname*{ess\,sup}_{a\in A} \mathbb{E}_{t_k,a}\big[\sup_{t\in[t_k,t_{k+1}]} |X_t - \bar X^\pi_{t_k}|^2\big]\ldots\big]\Big], \tag{2.2}
$$

where $\mathbb{E}_{t_k,a}[.]$ denotes the conditional expectation $\mathbb{E}[.|\mathcal{F}_{t_k}, I_{t_k} = a]$. We also denote by $\mathbb{E}_{t_k}[.]$ the conditional expectation $\mathbb{E}[.|\mathcal{F}_{t_k}]$. Since I_{t_k} is $\mathcal{F}_{t_k}$-measurable, and by the law of iterated conditional expectations, we notice that

$$
\mathbb{E}\Big[\sup_{t\in[t_k,t_{k+1}]} \big|X_t - \bar X^\pi_{t_k}\big|^2\Big] \leq \mathcal{E}^\pi_k(X), \quad k < n.
$$

Lemma 2.1 *We have*

$$
\max_{k<n} \mathcal{E}^\pi_k(X) \leq C_1|\pi|.
$$

Proof. From the definition of the Euler scheme, and under the growth linear condition in **(H1)**, we easily see that

$$
\mathbb{E}_{t_k}\Big[\big|\bar X^\pi_{t_{k+1}}\big|^2\Big] \leq C_1\big(1 + \big|\bar X^\pi_{t_k}\big|^2\big), \quad k < n. \tag{2.3}
$$

From the definition of the continuous-time Euler scheme, and by Burkholder-Davis-Gundy inequality, it is also clear that

$$
\mathbb{E}_{t_k}\Big[\sup_{t\in[t_k,t_{k+1}]} \big|\bar X^\pi_t - \bar X^\pi_{t_k}\big|^2\Big] \leq C_1(1 + \big|\bar X^\pi_{t_k}\big|^2)|\pi|, \quad k < n. \tag{2.4}
$$

We also have the standard estimate for the pure jump process I (recall that A is assumed to be compact and $\lambda(A) < \infty$):

$$
\mathbb{E}_{t_k}\Big[\sup_{t\in[t_k,t_{k+1}]} \big|I_s - I_{t_k}\big|^2\Big] \leq C_1|\pi|. \tag{2.5}
$$

Let us denote by $\Delta X_t = X_t - \bar{X}^\pi_t$, and apply Itô's formula to $|\Delta X_t|^2$ so that for all $t \in [t_k, t_{k+1}]$:

$$
\begin{aligned}
|\Delta X_t|^2 &= |\Delta X_{t_k}|^2 + \int_{t_k}^t 2\big(b(X_s, I_s) - b(\bar{X}^\pi_{t_k}, I_{t_k})\big).\Delta X_s + \big|\sigma(X_s, I_s) - \sigma(\bar{X}^\pi_{t_k}, I_{t_k})\big|^2 ds \\
&\quad + 2\int_{t_k}^t (\Delta X_s)'\big(\sigma(X_s, I_s) - \sigma(\bar{X}^\pi_{t_k}, I_{t_k})\big) dW_s \\
&\leq |\Delta X_{t_k}|^2 + C_1 \int_{t_k}^t |\Delta X_s|^2 + |\bar{X}^\pi_s - \bar{X}^\pi_{t_k}|^2 + |I_s - I_{t_k}|^2 ds \\
&\quad + 2\int_{t_k}^t (\Delta X_s)'\big(\sigma(X_s, I_s) - \sigma(\bar{X}^\pi_{t_k}, I_{t_k})\big) dW_s,
\end{aligned}
$$

from the Lipschitz condition on b, σ in **(H1)**. By taking conditional expectation in the above inequality, we then get:

$$
\begin{aligned}
\mathbb{E}_{t_k}\Big[|\Delta X_t|^2\Big] &\leq |\Delta X_{t_k}|^2 + C_1 \int_{t_k}^t \mathbb{E}_{t_k}\big[|\Delta X_s|^2 + |\bar{X}^\pi_s - \bar{X}^\pi_{t_k}|^2 + |I_s - I_{t_k}|^2\big] ds \\
&\leq |\Delta X_{t_k}|^2 + C_1(1 + \big|\bar{X}^\pi_{t_k}\big|^2)|\pi|^2 + C_1 \int_{t_k}^t \mathbb{E}_{t_k}\big[|\Delta X_s|^2\big] ds, \quad t \in [t_k, t_{k+1}],
\end{aligned}
$$

by (2.4)-(2.5). From Gronwall's lemma, we thus deduce that

$$
\mathbb{E}_{t_k}\Big[|\Delta X_{t_{k+1}}|^2\Big] \leq e^{C_1|\pi|}|\Delta X_{t_k}|^2 + C_1(1 + \big|\bar{X}^\pi_{t_k}\big|^2)|\pi|^2, \quad k < n. \tag{2.6}
$$

Since the right hand side of (2.6) does not depend on I_{t_k}, this shows that

$$
\operatorname*{ess\,sup}_{a \in A} \mathbb{E}_{t_k,a}\Big[|\Delta X_{t_{k+1}}|^2\Big] \leq e^{C_1|\pi|}|\Delta X_{t_k}|^2 + C_1(1 + \big|\bar{X}^\pi_{t_k}\big|^2)|\pi|^2.
$$

By taking conditional expectation w.r.t. $\mathcal{F}_{t_{k-1}}$ in the above inequality, using again estimate (2.6) together with (2.3) at step $k-1$, and iterating this backward procedure until the initial time $t_0 = 0$, we obtain:

$$
\begin{aligned}
&\mathbb{E}\Big[\operatorname*{ess\,sup}_{a \in A} \mathbb{E}_{t_1,a}\big[\ldots \operatorname*{ess\,sup}_{a \in A} \mathbb{E}_{t_k,a}\big[|\Delta X_{t_{k+1}}|^2\big] \ldots\big]\Big] \\
\leq\ & e^{C_1 n|\pi|}|\Delta X_0|^2 + C_1(1 + |X_0|^2)|\pi|^2 \frac{e^{C_1 n|\pi|} - 1}{e^{C_1|\pi|} - 1} \\
\leq\ & C_1|\pi|,
\end{aligned} \tag{2.7}
$$

since $\Delta X_0 = 0$ and $n|\pi|$ is bounded.

Moreover, the process X satisfies the standard conditional estimate similarly as for the Euler scheme:

$$
\begin{aligned}
\mathbb{E}_{t_k}\Big[\big|X_{t_{k+1}}\big|^2\Big] &\leq C_1\big(1 + \big|X_{t_k}\big|^2\big), \\
\mathbb{E}_{t_k}\Big[\sup_{t \in [t_k, t_{k+1}]} \big|X_t - X_{t_k}\big|^2\Big] &\leq C_1(1 + \big|X_{t_k}\big|^2)|\pi|, \quad k < n,
\end{aligned}
$$

from which we deduce by backward induction on the conditional expectations:

$$
\mathbb{E}\Big[\operatorname*{ess\,sup}_{a \in A} \mathbb{E}_{t_1,a}\big[\ldots \operatorname*{ess\,sup}_{a \in A} \mathbb{E}_{t_k,a}\big[\sup_{t \in [t_k, t_{k+1}]} \big|X_t - X_{t_k}\big|^2\big] \ldots\big]\Big] \leq C_1|\pi|. \tag{2.8}
$$

Finally, by writing that $\sup_{t\in[t_k,t_{k+1}]}|X_t - \bar{X}^\pi_{t_k}|^2 \leq 2\sup_{t\in[t_k,t_{k+1}]}|X_t - X_{t_k}|^2 + 2\Delta X_{t_k}$, taking successive condition expectations w.r.t to $\mathcal{F}_{t_\ell}$ and essential supremum over $I_{t_\ell} = a$, for ℓ going recursively from k to 0, we get:

$$
\begin{aligned}
\mathbb{E}_{t_k}\Big[\sup_{t\in[t_k,t_{k+1}]}|X_t - \bar{X}^\pi_{t_k}|^2\Big] &\leq 2\mathbb{E}\Big[\operatorname*{ess\,sup}_{a\in A}\mathbb{E}_{t_1,a}\big[\ldots\operatorname*{ess\,sup}_{a\in A}\mathbb{E}_{t_k,a}\big[\sup_{t\in[t_k,t_{k+1}]}\big|X_t - X_{t_k}\big|^2\big]\ldots\big]\Big] \\
&\quad + 2\mathbb{E}\Big[\operatorname*{ess\,sup}_{a\in A}\mathbb{E}_{t_1,a}\big[\ldots\operatorname*{ess\,sup}_{a\in A}\mathbb{E}_{t_{k-1},a}\big[|\Delta X_{t_k}|^2\big]\ldots\big]\Big] \\
&\leq C_1|\pi|,
\end{aligned}
$$

by (2.7)-(2.8), which ends the proof. □

3 Discretely jump-constrained BSDE

Given the forward regime switching process (X, I) defined in the previous section, we consider the minimal quadruple solution (Y, Z, U, K) to the BSDE with nonpositive jumps:

$$
\begin{cases}
Y_t = g(X_T) + \int_t^T f(X_s, I_s, Y_s, Z_s)ds + K_T - K_t \\
\qquad - \int_t^T Z_s dW_s - \int_t^T \int_A U_s(a)\tilde{\mu}(ds, da), \quad 0 \leq t \leq T. \\
U_t(a) \leq 0,
\end{cases}
\tag{3.1}
$$

By solution to (3.1), we mean a quadruple $(Y, Z, U, K) \in \mathcal{S}^2 \times L^2(W) \times L^2(\tilde{\mu}) \times \mathcal{K}^2$, where $\mathcal{S}^2$ is the space of càd-làg or càg-làd $\mathbb{F}$-progressively measurable processes Y satisfying $\|Y\|^2 := \mathbb{E}[\sup_{t\in[0,T]}|Y_t|^2] < \infty$, $L^2(W)$ is the space of $\mathbb{R}^d$-valued $\mathcal{P}$-measurable processes such that $\|Z\|^2_{L^2(W)} := \mathbb{E}[\int_0^T |Z_t|^2 dt] < \infty$, $L^2(\tilde{\mu})$ is the space of real-valued $\mathcal{P}\otimes\mathcal{B}(A)$-measurable processes U such that $\|U\|^2_{L^2(\tilde{\mu})} := \mathbb{E}[\int_0^T \int_A |U_t(a)|^2 \lambda(da)\, dt] < \infty$, and $\mathcal{K}^2$ is the subspace of $\mathcal{S}^2$ consisting of nondecreasing predictable processes such that $K_0 = 0$, $\mathbb{P}$-a.s., and the equation in (3.1) holds $\mathbb{P}$-a.s., while the nonpositive jump constraint holds on $\Omega \times [0, T] \times A$ a.e. with respect to the measure $d\mathbb{P} \otimes dt \otimes \lambda(da)$. By minimal solution to the BSDE (1.3), we mean a quadruple solution $(Y, Z, U, K) \in \mathcal{S}^2 \times L^2(W) \times L^2(\tilde{\mu}) \times \mathcal{K}^2$ such that for any other solution (Y', Z', U', K') to the same BSDE, we have $\mathbb{P}$-a.s.: $Y_t \leq Y'_t$, $t \in [0, T]$.

In the rest of this paper, we shall make the standing Lipschitz assumption on the functions $f : \mathbb{R}^d \times A \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ and $g : \mathbb{R}^d \to \mathbb{R}$.

(H2) There exists a constant L_2 such that

$$
|f(x, a, y, z) - f(x', a', y', z')| + |g(x) - g(x')| \leq L_2\big(|x - x'| + |a - a'| + |y - y'| + |z - z'|\big),
$$

for all $x, x' \in \mathbb{R}^d$, $y, y' \in \mathbb{R}$, $z, z' \in \mathbb{R}^d$, $a, a' \in A$. In the sequel, we shall denote by C a generic positive constant which depends only on L_1, L_2, T, (X_0, I_0) and $\lambda(A) < \infty$, and may vary from lines to lines.

Under **(H1)-(H2)**, it is proved in [13] the existence and uniqueness of a minimal solution (Y, Z, U, K) to (3.1). Moreover, the minimal solution Y is in the form

$$
Y_t = v(t, X_t), \quad 0 \leq t \leq T, \tag{3.2}
$$

where $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ is a viscosity solution with linear growth to the fully nonlinear HJB type equation:

$$\begin{cases} -\sup_{a\in A}\big[\mathcal{L}^a v + f(x,a,v,\sigma^\intercal(x,a)D_x v)\big] &=& 0, \text{ on } [0,T)\times\mathbb{R}^d, \\ v(T,x) &=& g, \text{ on } \mathbb{R}^d, \end{cases} \tag{3.3}$$

where

$$\mathcal{L}^a v \;=\; \frac{\partial v}{\partial t} + b(x,a).D_x v + \frac{1}{2}\mathrm{tr}(\sigma\sigma^\intercal(x,a)D_x^2 v).$$

We shall make the standing assumption that comparison principle holds for (3.3).

(HC) Let $\bar{w}$ (resp. $\underline{w}$) be a lower-semicontinuous (resp. upper-semicontinuous) viscosity supersolution (resp. subsolution) with linear growth condition to (3.3). Then, $\bar{w} \geq \underline{w}$.

When f does not depend on y, z, i.e. (3.3) is the usual HJB equation for a stochastic control problem, Assumption **(HC)** holds true, see [11] or [19]. In the general case, we refer to [9] for sufficient conditions to comparison principles. Under **(HC)**, the function v in (3.2) is the unique viscosity solution to (3.3), and is in particular continuous. Actually, we have the standard Hölder and Lipschitz property (see Appendix in [17] or [4]):

$$|v(t,x) - v(t',x')| \;\leq\; C\big(|t-t'|^{\frac{1}{2}} + |x-x'|\big), \quad (t,t') \in [0,T], x, x' \in \mathbb{R}^d. \tag{3.4}$$

This implies that the process Y is continuous, and thus the jump component $U = 0$. In the sequel, we shall focus on the approximation of the remaining components Y and Z of the minimal solution to (3.1).

We introduce in this section discretely jump-constrained BSDE. The nonpositive jump constraint operates only at the times of the grid $\pi = \{t_0 = 0 < t_1 < \ldots < t_n = T\}$ of $[0,T]$, and we look for a quadruple $(Y^\pi, \mathcal{Y}^\pi, \mathcal{Z}^\pi, \mathcal{U}^\pi) \in \mathcal{S}^2 \times \mathcal{S}^2 \times L^2(W) \times L^2(\tilde{\mu})$ satisfying:

$$Y_T^\pi \;=\; \mathcal{Y}_T^\pi \;=\; g(X_T) \tag{3.5}$$

and

$$\begin{aligned} \mathcal{Y}_t^\pi \;&=\; Y_{t_{k+1}}^\pi + \int_t^{t_{k+1}} f(X_s, I_s, \mathcal{Y}_s^\pi, \mathcal{Z}_s^\pi)ds \qquad (3.6) \\ &\quad - \int_t^{t_{k+1}} \mathcal{Z}_s^\pi dW_s - \int_t^{t_{k+1}}\int_A \mathcal{U}_s^\pi(a)\tilde{\mu}(ds,da)\;, \\ Y_t^\pi \;&=\; \mathcal{Y}_t^\pi \mathbb{1}_{(t_k,t_{k+1})}(t) + \operatorname*{ess\,sup}_{a\in A} \mathbb{E}\Big[\mathcal{Y}_t^\pi \Big| X_t, I_t \;=\; a\Big]\mathbb{1}_{\{t_k\}}(t)\;, \qquad (3.7) \end{aligned}$$

for all $t \in [t_k, t_{k+1})$ and all $0 \leq k \leq n-1$.

Notice that at each time t_k of the grid, the condition is not known a priori to be square integrable since it involves a supremum over A, and the well-posedness of the BSDE (3.5)-(3.6)-(3.7) is not a direct and standard issue. We shall use a PDE approach for proving the existence and uniqueness of a solution. Let us consider the system of integro-partial differential equations (IPDEs) for the functions v^π and ϑ^π defined recursively on $[0,T] \times \mathbb{R}^d \times A$ by:

• A terminal condition for v^π and ϑ^π:

$$v^\pi(T,x,a) \;=\; \vartheta^\pi(T,x,a) \;=\; g(x)\;, \quad (x,a)\in\mathbb{R}^d\times A\;, \tag{3.8}$$

• A sequence of IPDEs for ϑ^π

$$\left\{\begin{array}{rl} -\mathcal{L}^a\vartheta^\pi - f\big(x,a,\vartheta^\pi,\sigma^\intercal(x,a)D_x\vartheta^\pi\big) & \\ -\int_A\big(\vartheta^\pi(t,x,a')-\vartheta^\pi(t,x,a)\big)\lambda(da') \;=\; 0, & (t,x,a)\in[t_k,t_{k+1})\times\mathbb{R}^d\times A,\\ \vartheta^\pi(t_{k+1}^-,x,a) \;=\; \sup_{a'\in A}\vartheta^\pi(t_{k+1},x,a') & (x,a)\in\mathbb{R}^d\times A \end{array}\right. \tag{3.9}$$

for $k=0\ldots,n-1$,

• the relation between v^π and ϑ^π:

$$v^\pi(t,x,a) \;=\; \vartheta^\pi(t,x,a)\mathbb{1}_{(t_k,t_{k+1})}(t)+\sup_{a'\in A}\vartheta^\pi(t,x,a')\mathbb{1}_{\{t_k\}}(t)\;, \tag{3.10}$$

for all $t\in[t_k,t_{k+1})$ and $k=0\ldots,n-1$. The rest of this section is devoted to the proof of existence and uniqueness of a solution to (3.8)-(3.9)-(3.10), together with some uniform Lipschitz properties, and its connection to the discretely jump-constrained BSDE (3.5)-(3.6)-(3.7).

For any L-Lipschitz continuous function φ on $\mathbb{R}^d\times A$, and $k\le n-1$, we denote:

$$\mathbb{T}_\pi^k[\varphi](t,x,a) \;:=\; w(t,x,a), \quad (t,x,a)\in[t_k,t_{k+1})\times\mathbb{R}^d\times A, \tag{3.11}$$

where w is the unique continuous viscosity solution on $[t_k,t_{k+1}]\times\mathbb{R}^d\times A$ with linear growth condition in x to the integro partial differential equation (IPDE):

$$\left\{\begin{array}{rl} -\mathcal{L}^a w - f(x,a,w,\sigma^\intercal D_x w) & \\ -\int_A\big(w(t,x,a')-w(t,x,a)\big)\lambda(da') \;=\; 0, & (t,x,a)\in[t_k,t_{k+1})\times\mathbb{R}^d\times A,\\ w(t_{k+1}^-,x,a) \;=\; \varphi(x,a), & (x,a)\in\mathbb{R}^d\times A\;, \end{array}\right. \tag{3.12}$$

and we extend by continuity $\mathbb{T}_\pi^k[\varphi](t_{k+1},x,a)=\varphi(x,a)$. The existence and uniqueness of such a solution w to the semi linear IPDE (3.12), and its nonlinear Feynman-Kac representation in terms of BSDE with jumps, is obtained e.g. from Theorems 3.4 and 3.5 in [3].

Lemma 3.1 *There exists a constant C such that for any L-Lipschitz continuous function φ on $\mathbb{R}^d\times A$, and $k\le n-1$, we have*

$$|\mathbb{T}_\pi^k[\varphi](t,x,a)-\mathbb{T}_\pi^k[\varphi](t,x',a')| \;\le\; \max(L,1)\sqrt{1+|\pi|}e^{C|\pi|}(|x-x'|+|a-a'|)\;,$$

for all $t\in[t_k,t_{k+1})$, and $(x,a),(x',a')\in\mathbb{R}^d\times A$.

Proof. Fix $t\in[t_k,t_{k+1})$, $k\le n-1$, $(x,a),(x',a')\in\mathbb{R}^d\times A$, and φ an L-Lipschitz continuous function on $\mathbb{R}^d\times A$. Let $(Y^\varphi,Z^\varphi,U^\varphi)$ and $(Y^{\varphi,'},Z^{\varphi,'},U^{\varphi,'})$ be the solutions on $[t,t_{k+1}]$ to the BSDEs

$$
\begin{aligned}
Y_s^{\varphi} &= \varphi(X_{t_{k+1}}^{t,x,a}, I_{t_{k+1}}^{t,a}) + \int_s^{t_{k+1}} f(X_r^{t,x,a}, I_r^{t,a}, Y_r^{\varphi}, Z_r^{\varphi})dr \\
&\quad - \int_s^{t_{k+1}} Z_r^{\varphi} dW_r - \int_s^{t_{k+1}} \int_A U_r^{\varphi}(e)\tilde{\mu}(dr, de), \qquad t \le s \le t_{k+1}, \\
Y_s^{\varphi,'} &= \varphi(X_{t_{k+1}}^{t,x',a'}, I_{t_{k+1}}^{t,a'}) + \int_s^{t_{k+1}} f(X_r^{t,x',a'}, I_r^{t,a'}, Y_r^{\varphi,'}, Z_r^{\varphi,'})dr \\
&\quad - \int_s^{t_{k+1}} Z_r^{\varphi,'} dW_r - \int_s^{t_{k+1}} \int_A U_r^{\varphi,'}(e)\tilde{\mu}(dr, de), \qquad t \le s \le t_{k+1}
\end{aligned}
$$

From Theorems 3.4 and 3.5 in [3], we have the identification:

$$
Y_t^{\varphi} = \mathbb{T}_\pi^k[\varphi](t,x,a) \quad \text{and} \quad Y_t^{\varphi,'} = \mathbb{T}_\pi^k[\varphi](t,x',a') . \tag{3.13}
$$

We now estimate the difference between the processes $Y^{\varphi}$and $Y^{\varphi,'}$, and set $\delta Y^{\varphi} = Y^{\varphi} - Y^{\varphi,'}$, $\delta Z^{\varphi} = Z^{\varphi} - Z^{\varphi,'}$, $\delta X = X^{t,x,a} - X^{t,x',a'}$, $\delta I = I^{t,a} - I^{t,a'}$. By Itô's formula, the Lipschitz condition of f and φ, and Young inequality, we have

$$
\begin{aligned}
\mathbb{E}\Big[|\delta Y_s^{\varphi}|^2\Big] + \mathbb{E}\Big[\int_s^{t_{k+1}} |\delta Z_s^{\varphi}|^2 ds\Big] &\le L^2 \mathbb{E}\Big[|\delta X_T|^2 + |\delta I_T|^2\Big] + C\int_s^{t_{k+1}} \mathbb{E}\Big[|\delta Y_r^{\varphi}|^2\Big] dr \\
&\quad + \frac{1}{2}\mathbb{E}\Big[\int_s^{t_{k+1}} \big(|\delta X_r|^2 + |\delta I_r|^2 + |\delta Z_r^{\varphi}|^2\big) dr\Big],
\end{aligned}
$$

for any $s \in [t, t_{k+1}]$. Now, from classical estimates on jump-diffusion processes we have

$$
\sup_{r\in[t,t_{k+1}]} \mathbb{E}\Big[|\delta X_r|^2 + |\delta I_r|^2\Big] \le e^{C|\pi|}\big(|x-x'|^2 + |a-a'|^2\big),
$$

and thus:

$$
\mathbb{E}\Big[|\delta Y_s^{\varphi}|^2\Big] \le (L^2 + |\pi|)e^{C|\pi|}\big(|x-x'|^2 + |a-a'|^2\big) + C\int_s^{t_{k+1}} \mathbb{E}\Big[|\delta Y_r^{\varphi}|^2\Big] dr ,
$$

for all $s \in [t, t_{k+1}]$. By Gronwall's Lemma, this yields

$$
\sup_{s\in[t,t_{k+1}]} \mathbb{E}\Big[|\delta Y_s^{\varphi}|^2\Big] \le (L^2 + |\pi|)e^{2C|\pi|}\big(|x-x'|^2 + |a-a'|^2\big),
$$

which proves the required result from the identification (3.13):

$$
\begin{aligned}
|\mathbb{T}_\pi^k[\varphi](t,x,a) - \mathbb{T}_\pi^k[\varphi](t,x',a')| &\le \sqrt{L^2 + |\pi|}e^{C|\pi|}(|x-x'| + |a-a'|) \\
&\le \max(L,1)\sqrt{1+|\pi|}e^{C|\pi|}(|x-x'| + |a-a'|).
\end{aligned}
$$

□

Proposition 3.1 *There exists a unique viscosity solution* ϑ^π *with linear growth condition to the IPDE* (3.8)-(3.9)*, and this solution satisfies:*

$$
\begin{aligned}
&|\vartheta^\pi(t,x,a) - \vartheta^\pi(t,x',a')| \\
\le\ &\max(L_2, 1)\sqrt{\Big(e^{2C|\pi|}(1+|\pi|)\Big)^{n-k}}\big(|x-x'| + |a-a'|\big),
\end{aligned} \tag{3.14}
$$

for all $k = 0, \ldots, n-1$, $t \in [t_k, t_{k+1})$, $(x,a), (x',a') \in \mathbb{R}^d \times A$.

Proof. We prove by a backward induction on k that the IPDE (3.8)-(3.9) admits a unique solution on $[t_k, T] \times \mathbb{R}^d \times A$, which satisfies (3.14).

• For $k = n-1$, we directly get the existence and uniqueness of ϑ^π on $[t_{n-1}, T] \times \mathbb{R}^d \times A$ from Theorems 3.4 and 3.5 in [3], and we have $\vartheta^\pi = \mathbb{T}_\pi^{n-1}[g]$ on $[t_{n-1}, T) \times \mathbb{R}^d \times A$. Moreover, we also get by Lemma 3.1:

$$|\vartheta^\pi(t,x,a) - \vartheta^\pi(t,x',a')| \quad \leq \quad \max(L_2, 1)\sqrt{e^{2C|\pi|}(1+|\pi|)}\big(|x-x'| + |a-a'|\big)$$

for all $t \in [t_{n-1}, t_n)$, $(x,a), (x',a') \in \mathbb{R}^d \times A$.

• Suppose that the result holds true at step $k+1$ i.e. there exists a unique function ϑ^π on $[t_{k+1}, T] \times \mathbb{R}^d \times A$ with linear growth and satisfying (3.8)-(3.9) and (3.14). It remains to prove that ϑ^π is uniquely determined by (3.9) on $[t_k, t_{k+1}) \times \mathbb{R}^d \times A$ and that it satisfies (3.14) on $[t_k, t_{k+1}) \times \mathbb{R}^d \times A$. Since ϑ^π satisfies (3.14) at time t_{k+1}, we deduce that the function

$$\psi_{k+1}(x) \quad := \quad \sup_{a \in A} \vartheta^\pi(t_{k+1}, x, a), \quad x \in \mathbb{R}^d,$$

is also Lipschitz continuous, and satisfies by the induction hypothesis:

$$|\psi_{k+1}(x) - \psi_{k+1}(x')| \quad \leq \quad \max(L_2, 1)\sqrt{\Big(e^{2C|\pi|}(1+|\pi|)\Big)^{n-k-1}}|x-x'|, \tag{3.15}$$

for all $x, x' \in \mathbb{R}^d$. Under **(H1)** and **(H2)**, we can apply Theorems 3.4 and 3.5 in [3], and we get that ϑ^π is the unique viscosity solution with linear growth to (3.9) on $[t_k, t_{k+1}) \times \mathbb{R}^d \times A$, with $\vartheta^\pi = \mathbb{T}_\pi^k[\psi_{k+1}]$. Thus it exists and is unique on $[t_k, T] \times \mathbb{R}^d \times A$. From Lemma 3.1 and (3.15), we then get

$$\begin{aligned}
|\vartheta^\pi(t,x,a) - \vartheta^\pi(t,x',a')| \quad &= \quad |\mathbb{T}_\pi^k[\psi_{k+1}](t,x,a) - \mathbb{T}_\pi^k[\psi_{k+1}](t,x',a')| \\
&\leq \quad \max(L_2, 1)\sqrt{\Big(e^{2C|\pi|}(1+|\pi|)\Big)^{n-k-1}} \\
&\qquad .\sqrt{(1+|\pi|)e^{2C|\pi|}}\big(|x-x'| + |a-a'|\big) \\
&\leq \quad \max(L_2, 1)\sqrt{\Big(e^{2C|\pi|}(1+|\pi|)\Big)^{n-k}}\big(|x-x'| + |a-a'|\big)
\end{aligned}$$

for any $t \in [t_k, t_{k+1})$ and $(x,a), (x',a') \in \mathbb{R}^d \times A$, which proves the required induction inequality at step k. □

Remark 3.1 The function $a \to \vartheta^\pi(t,x,.)$ is continuous on A, for each (t,x), and so the function v^π is well-defined by (3.10). Moreover, the function ϑ^π may be written recursively as:

$$\begin{cases} \vartheta^\pi(T,.,.) &= \quad g \qquad\qquad\qquad\quad \text{on} \ \ \mathbb{R}^d \times A, \\ \vartheta^\pi &= \quad \mathbb{T}_\pi^k[v^\pi(t_{k+1},.)], \quad \text{on} \ \ [t_k, t_{k+1}) \times \mathbb{R}^d \times A, \end{cases} \tag{3.16}$$

for $k = 0, \ldots, n-1$. In particular, ϑ^π is continuous on $(t_k, t_{k+1}) \times \mathbb{R}^d \times A$, $k \leq n-1$. □

As a consequence of the above proposition, we obtain the uniform Lipschitz property of ϑ^π and v^π, with a Lipschitz constant independent of π.

Corollary 3.1 *There exists a constant C (independent of $|\pi|$) such that*

$$|\vartheta^\pi(t,x,a) - \vartheta^\pi(t,x',a')| + |v^\pi(t,x,a) - v^\pi(t,x',a')| \quad \leq \quad C\big(|x-x'| + |a-a'|\big),$$

for all $t \in [0,T]$, $x, x' \in \mathbb{R}^d$, $a, a' \in \mathbb{R}^d$.

Proof. Recalling that $n|\pi|$ is bounded, we see that the sequence appearing in (3.14): $\Big(\big(e^{2C|\pi|}(1+|\pi|)\big)^{n-k}\Big)_{0\leq k\leq n-1}$ is bounded uniformly in $|\pi|$ (or n), which shows the required Lipschitz property of ϑ^π. Since A is assumed to be compact, this shows in particular that the function v^π defined by the relation (3.10) is well-defined and finite. Moreover, by noting that

$$|\sup_{a\in A}\vartheta^\pi(t,x,a) - \sup_{a\in A}\vartheta^\pi(t,x',a)| \quad \leq \quad \sup_{a\in A}|\vartheta^\pi(t,x,a) - \vartheta^\pi(t,x',a)|$$

for all $(t,x) \in [0,T]\times\mathbb{R}^d$, we also obtain the required Lipschitz property for v^π. $\square$

We now turn to the existence of a solution to the discretely jump-constrained BSDE.

Proposition 3.2 *The BSDE* (3.5)-(3.6)-(3.7) *admits a unique solution* $(Y^\pi, \mathcal{Y}^\pi, \mathcal{Z}^\pi, \mathcal{U}^\pi)$ *in $\mathcal{S}^2 \times \mathcal{S}^2 \times L^2(W) \times L^2(\tilde{\mu})$. Moreover we have*

$$\mathcal{Y}^\pi_t \quad = \quad \vartheta^\pi(t,X_t,I_t), \quad \text{and} \quad Y^\pi_t \quad = \quad v^\pi(t,X_t,I_t) \tag{3.17}$$

for all $t \in [0,T]$.

Proof. We prove by backward induction on k that $(Y^\pi, \mathcal{Y}^\pi, \mathcal{Z}^\pi, \mathcal{U}^\pi)$ is well defined and satisfies (3.17) on $[t_k, T]$.

• Suppose that $k = n-1$. From Corollary 2.3 in [3], we know that $(\mathcal{Y}^\pi, \mathcal{Z}^\pi, \mathcal{U}^\pi)$, exists and is unique on $[t_{n-1}, T]$. Moreover, from Theorems 3.4 and 3.5 in [3], we get $\mathcal{Y}^\pi_t = \mathbb{T}^k_\pi[g](t,X_t,I_t) = \vartheta^\pi(t,X_t,I_t)$ on $[t_{n-1},T]$. By (3.7), we then have for all $t \in [t_{n-1}, T)$:

$$\begin{aligned} Y^\pi_t &= \mathbb{1}_{(t_{n-1},T)}(t)\,\vartheta^\pi(t,X_t,I_t) + \mathbb{1}_{t_{n-1}}(t)\,\operatorname*{ess\,sup}_{a\in A}\vartheta^\pi(t,X_t,a) \\ &= \mathbb{1}_{(t_{n-1},T)}(t)\,\vartheta^\pi(t,X_t,I_t) + \mathbb{1}_{t_{n-1}}(t)\,\sup_{a\in A}\vartheta^\pi(t,X_t,a) \;=\; v^\pi(t,X_t,I_t), \end{aligned}$$

since the essential supremum and supremum coincide by continuity of $a \to \vartheta^\pi(t,X_t,a)$ on the compact set A.

• Suppose that the result holds true for some $k \leq n-1$. Then, we see that $(\mathcal{Y}^\pi, \mathcal{Z}^\pi, \mathcal{U}^\pi)$ is defined on $[t_{k-1}, t_k)$ as the solution to a BSDE driven by W and $\tilde{\mu}$ with a terminal condition $v^\pi(t_k, X_{t_k})$. Since v^π satisfies a linear growth condition, we know again by Corollary 2.3 in [3] that $(\mathcal{Y}^\pi, \mathcal{Z}^\pi, \mathcal{U}^\pi)$, thus also Y^π, exists and is unique on $[t_{k-1}, t_k)$. Moreover, using again Theorems 3.4 and 3.5 in [3], we get (3.17) on $[t_{k-1}, t_k)$. $\square$

We end this section with a conditional regularity result for the discretely jump-constrained BSDE.

Proposition 3.3 *There exists some constant C such that*

$$\sup_{t\in[t_k,t_{k+1})} \mathbb{E}_{t_k}\big[|\mathcal{Y}^\pi_t - \mathcal{Y}^\pi_{t_k}|^2\big] + \sup_{t\in(t_k,t_{k+1}]} \mathbb{E}_{t_k}\big[|Y^\pi_t - Y^\pi_{t_{k+1}}|^2\big] \;\; \leq \;\; C(1+|X_{t_k}|^2)|\pi|,$$

for all $k = 0,\ldots,n-1$.

Proof. Fix $k \leq n-1$. By Itô's formula, we have for all $t \in [t_k, t_{k+1})$:

$$\begin{aligned}
\mathbb{E}_{t_k}\big[|\mathcal{Y}^\pi_t - \mathcal{Y}^\pi_{t_k}|^2\big] &= 2\mathbb{E}_{t_k}\Big[\int_{t_k}^t f(X_s,I_s,\mathcal{Y}^\pi_s,\mathcal{Z}^\pi_s)(\mathcal{Y}^\pi_{t_k} - \mathcal{Y}^\pi_s)ds\Big] \\
&\quad + \mathbb{E}_{t_k}\Big[\int_{t_k}^t |\mathcal{Z}^\pi_s|^2\Big] + \mathbb{E}_{t_k}\Big[\int_{t_k}^t \int_A |\mathcal{U}^\pi_s(a)|^2\lambda(da)ds\Big] \\
&\leq \mathbb{E}_{t_k}\Big[\int_{t_k}^t |\mathcal{Y}^\pi_s - \mathcal{Y}^\pi_{t_k}|^2\Big] + C|\pi|\Big(1 + \mathbb{E}_{t_k}\Big[\sup_{s\in[t_k,t_{k+1}]}|X_s|^2\Big]\Big) \\
&\quad + C|\pi|\mathbb{E}_{t_k}\Big[\sup_{s\in[t_k,t_{k+1}]}\Big(|\mathcal{Y}^\pi_s|^2 + |\mathcal{Z}^\pi_s|^2 + \int_A |\mathcal{U}^\pi_s(a)|^2\lambda(da)\Big)\Big],
\end{aligned}$$

by the linear growth condition on f (recall also that A is compact), and Young inequality. Now, by standard estimate for X under growth linear condition on b and σ, we have:

$$\mathbb{E}_{t_k}\Big[\sup_{s\in[t_k,t_{k+1}]}|X_s|^2\Big] \;\; \leq \;\; C(1+|X_{t_k}|^2). \tag{3.18}$$

We also know from Proposition 4.2 in [7], under **(H1)** and **(H2)**, that there exists a constant C depending only on the Lipschitz constants of b, σ f and $v^\pi(t_{k+1},.)$ (which does not depend on π by Corollary 3.1), such that

$$\mathbb{E}_{t_k}\Big[\sup_{s\in[t_k,t_{k+1}]}\Big(|\mathcal{Y}^\pi_s|^2 + |\mathcal{Z}^\pi_s|^2 + \int_A |\mathcal{U}^\pi_s(a)|^2\lambda(da)\Big)\Big] \;\; \leq \;\; C(1+|X_{t_k}|^2). \tag{3.19}$$

We deduce that

$$\mathbb{E}_{t_k}\big[|\mathcal{Y}^\pi_t - \mathcal{Y}^\pi_{t_k}|^2\big] \;\; \leq \;\; \mathbb{E}_{t_k}\Big[\int_{t_k}^t |\mathcal{Y}^\pi_s - \mathcal{Y}^\pi_{t_k}|^2\Big] + C|\pi|(1+|X_{t_k}|^2),$$

and we conclude for the regularity of $\mathcal{Y}^\pi$ by Gronwall's lemma. Finally, from the definition (3.6)-(3.7) of Y^π and $\mathcal{Y}^\pi$, Itô isometry for stochastic integrals, and growth linear condition on f, we have for all $t \in (t_k, t_{k+1})$:

$$\begin{aligned}
\mathbb{E}_{t_k}\big[|Y^\pi_t - Y^\pi_{t_{k+1}}|^2\big] &= \mathbb{E}_{t_k}\big[|\mathcal{Y}^\pi_t - Y^\pi_{t_{k+1}}|^2\big] \\
&\leq 3\mathbb{E}_{t_k}\Big[\int_{t_k}^{t_{k+1}}\Big(|f(X_s,I_s,\mathcal{Y}^\pi_s,\mathcal{Z}^\pi_s)|^2 + |\mathcal{Z}^\pi_s|^2 + \int_A |\mathcal{U}^\pi_s(a)|^2\lambda(da)\Big)ds\Big] \\
&\leq C|\pi|\mathbb{E}_{t_k}\Big[1 + \sup_{s\in[t_k,t_{k+1}]}\Big(|X_s|^2 + |\mathcal{Y}^\pi_s|^2 + |\mathcal{Z}^\pi_s|^2 + \int_A |\mathcal{U}^\pi_s(a)|^2\lambda(da)\Big)\Big] \\
&\leq C|\pi|(1+|X_{t_k}|^2),
\end{aligned}$$

where we used again (3.18) and (3.19). This ends the proof. □

4 Convergence of discretely jump-constrained BSDE

This section is devoted to the convergence of the discretely jump-constrained BSDE towards the minimal solution to the BSDE with nonpositive jump.

4.1 Convergence result

Lemma 4.1 *We have the following assertions:*
1) The familly $(\vartheta^\pi)_\pi$ is nondecreasing and upper bounded by v: for any grids π and π' such that $\pi \subset \pi'$, we have

$$\vartheta^\pi(t,x,a) \;\; \leq \;\; \vartheta^{\pi'}(t,x,a) \; \leq \; v(t,x) \; , \quad (t,x,a) \in [0,T] \times \mathbb{R}^d \times A \; .$$

2) The familly $(\vartheta^\pi)_\pi$ satisfies a uniform linear growth condition: there exists a constant C such that

$$|\vartheta^\pi(t,x,a)| \;\; \leq \;\; C(1+|x|),$$

for any $(t,x,a) \in [0,T] \times \mathbb{R}^d \times A$ and any grid π.

Proof. 1) Let us first prove that $\vartheta^\pi \leq v$. Since v is a (continuous) viscosity solution to the HJB equation (3.3), and v does not depend on a, we see that v is a viscosity supersolution to the IPDE in (3.9) satisfied by ϑ^π on each interval $[t_k, t_{k+1})$. Now, since $v(T,x) = \vartheta^\pi(T,x,a)$, we deduce by comparison principle for this IPDE (see e.g. Theorem 3.4 in [3]) on $[t_{n-1}, T) \times \mathbb{R}^d \times A$ that $v(t,x) \geq \vartheta^\pi(t,x,a)$ for all $t \in [t_{n-1}, T]$, $(x,a) \in \mathbb{R}^d \times A$. In particular, $v(t_{n-1}^-, x) = v(t_{n-1}, x) \geq \sup_{a \in A} \vartheta^\pi(t_{n-1}, x, a) = \vartheta^\pi(t_{n-1}^-, x, a)$. Again, by comparison principle for the IPDE (3.9) on $[t_{n-2}, t_{n-1}) \times \mathbb{R}^d \times A$, it follows that $v(t,x)$ $\geq \vartheta^\pi(t,x,a)$ for all $t \in [t_{n-2}, t_{n-1}]$, $(x,a) \in \mathbb{R}^d \times A$. By backward induction on time, we conclude that $v \geq \vartheta^\pi$ on $[0,T] \times \mathbb{R}^d \times A$.

Let us next consider two partitions $\pi = (t_k)_{0 \leq k \leq n}$ and $\pi' = (t'_k)_{0 \leq k \leq n'}$ of $[0,T]$ with π $\subset \pi'$, and denote by $m = \max\{k \leq n' : t'_m \notin \pi\}$. Thus, all the points of the grid π and π' coincide after time t'_m, and since ϑ^π and $\vartheta^{\pi'}$ are viscosity solution to the same IPDE (3.9) starting from the same terminal data g, we deduce by uniqueness that $\vartheta^\pi = \vartheta^{\pi'}$ on $[t'_m, T] \times \mathbb{R}^d \times A$. Then, we have $\vartheta^{\pi'}(t'^-_m, x, a) = \sup_{a \in A} \vartheta^\pi(t'_m, x, a) = \sup_{a \in A} \vartheta^\pi(t'_m, x, a)$ $\geq \vartheta^\pi(t^-_m, x, a)$ since ϑ^π is continuous outside of the points of the grid π (recall Remark 3.1). Now, since ϑ^π and $\vartheta^{\pi'}$ are viscosity solution to the same IPDE (3.9) on $[t'_{m-1}, t_m)$, we deduce by comparison principle that $\vartheta^{\pi'} \geq \vartheta^\pi$ on $[t'_{m-1}, t'_m] \times \mathbb{R}^d \times A$. Proceeding by backward induction, we conclude that $\vartheta^{\pi'} \geq \vartheta^\pi$ on $[0,T] \times \mathbb{R}^d \times A$.

2) Denote by $\pi_0 = \{t_0 = 0, t_1 = T\}$ the trivial grid of $[0,T]$. Since $\vartheta^{\pi_0} \leq \vartheta^\pi \leq v$ and ϑ^{π_0} and v satisfy a linear growth condition, we get (recall that A is compact):

$$|\vartheta^\pi(t,x,a)| \;\; \leq \;\; |\vartheta^{\pi_0}(t,x,a)| + |v(t,x)| \; \leq \; C(1+|x|),$$

for any $(t,x,a) \in [0,T] \times \mathbb{R}^d \times A$ and any grid π. □

In the sequel, we denote by ϑ the increasing limit of the sequence $(\vartheta^\pi)_\pi$ when the grid increases by becoming finer, i.e. its modulus $|\pi|$ goes to zero. The next result shows that ϑ does not depend on the variable a in A.

Proposition 4.1 *The function ϑ is l.s.c. and does not depend on the variable $a \in A$:*

$$\vartheta(t,x,a) \quad = \quad \vartheta(t,x,a') \ , \quad (t,x) \in [0,T] \times \mathbb{R}^d, \ a, a' \in A \ .$$

To prove this result we use the following lemma. Observe by definition (3.10) of v^π that the function v^π does not depend on a on the grid times π, and we shall denote by misuse of notation: $v^\pi(t_k, x)$, for $k \leq n$, $x \in \mathbb{R}^d$.

Lemma 4.2 *There exists a constant C (not depending on π) such that*

$$|\vartheta^\pi(t,x,a) - v^\pi(t_{k+1},x)| \quad \leq \quad C(1+|x|)|\pi|^{\frac{1}{2}}$$

for all $k = 0, \ldots, n-1$, $t \in [t_k, t_{k+1})$, $(x,a) \in \mathbb{R}^d \times A$.

Proof. Fix $k = 0, \ldots, n-1$, $t \in [t_k, t_{k+1})$ and $(x,a) \in \mathbb{R}^d \times A$. Let $(\tilde{\mathcal{Y}}, \tilde{\mathcal{Z}}, \tilde{\mathcal{U}})$ be the solution to the BSDE

$$\begin{aligned} \tilde{\mathcal{Y}}_s \quad &= \quad v^\pi(t_{k+1}, X^{t,x,a}_{t_{k+1}}) + \int_s^{t_{k+1}} f(X^{t,x,a}_s, I^{t,a}_s, \tilde{\mathcal{Y}}_s, \tilde{\mathcal{Z}}_s) ds \\ & \quad - \int_s^{t_{k+1}} \tilde{\mathcal{Z}}_s dW_s - \int_s^{t_{k+1}} \int_A \tilde{\mathcal{U}}_s(a') \tilde{\mu}(ds, da') \ , \qquad s \in [t, t_{k+1}] \ . \end{aligned}$$

From Proposition 3.2, Markov property and uniqueness of a solution to the BSDE (3.5)-(3.6)-(3.7) we have: $\tilde{\mathcal{Y}}_s = \vartheta^\pi(s, X^{t,x,a}_s, I^{t,a}_s)$, for $s \in [t, t_{k+1}]$, and so:

$$\begin{aligned} |\vartheta^\pi(t,x,a) - v^\pi(t_{k+1},x)| \quad &= \quad \big|\tilde{\mathcal{Y}}_t - v^\pi(t_{k+1},x)\big| \\ &\leq \quad \mathbb{E}|v^\pi(t_{k+1}, X^{t,x,a}_{t_{k+1}}) - v^\pi(t_{k+1},x)| \\ & \qquad + \mathbb{E}\Big[\int_t^{t_{k+1}} \big|f(X^{t,x,a}_s, I^{t,a}_s, \tilde{\mathcal{Y}}_s, \tilde{\mathcal{Z}}_s)\big| ds\Big]. \end{aligned} \tag{4.1}$$

From Corollary 3.1, we have

$$\mathbb{E}|v^\pi(t_{k+1}, X^{t,x,a}_{t_{k+1}}) - v^\pi(t_{k+1},x)| \quad \leq \quad C\sqrt{\mathbb{E}[|X^{t,x,a}_{t_{k+1}} - x|^2]} \ \leq \ C\sqrt{|\pi|} \ . \tag{4.2}$$

Moreover, by the growth linear condition on f in **(H2)**, and on ϑ^π in Lemma 4.1, we have

$$\mathbb{E}\Big[\int_t^{t_{k+1}} \big|f(X_s, I_s, \tilde{\mathcal{Y}}_s, \tilde{\mathcal{Z}}_s)\big| ds\Big] \quad \leq \quad C\mathbb{E}\Big[\int_t^{t_{k+1}} \Big(1 + |X^{t,x,a}_s| + |\tilde{\mathcal{Z}}_s|\Big) ds\Big].$$

By classical estimates, we have

$$\sup_{s \in [t,T]} \mathbb{E}\big[|X^{t,x,a}_s|^2\big] \quad \leq \quad C(1+|x|^2).$$

Moreover, under **(H1)** and **(H2)**, we know from Proposition 4.2 in [7] that there exists a constant C depending only on the Lipschitz constants of b, σ f and $v^\pi(t_{k+1}, .)$ such that

$$\mathbb{E}\big[\sup_{s \in [t_k, t_{k+1}]} |\tilde{\mathcal{Z}}_s|^2\big] \quad \leq \quad C(1+|x|^2).$$

This proves that

$$\mathbb{E}\Big[\int_t^{t_{k+1}} \big|f(X_s, I_s, \tilde{\mathcal{Y}}_s, \tilde{\mathcal{Z}}_s)\big|ds\Big] \leq C(1+|x|)|\pi| \ .$$

Combining this last estimate with (4.1) and (4.2), we get the result □

Proof of Proposition 4.1. The function ϑ is l.s.c. as the supremum of the l.s.c. functions ϑ^π. Fix $(t,x) \in [0,T) \times \mathbb{R}^d$ and $a, a' \in A$. Let $(\pi^p)_p$ be a sequence of subdivisions of $[0,T]$ such that $|\pi^p| \downarrow 0$ as $p \uparrow \infty$. We define the sequence $(t_p)_p$ of $[0,T]$ by

$$t_p = \min\big\{s \in \pi^p \ : \ s > t\big\} \ , \quad p \geq 0 \ .$$

Since $|\pi^p| \to 0$ as $p \to \infty$ we get $t_p \to t$ as $p \to +\infty$. We then have from the previous lemma:

$$\begin{aligned} |\vartheta^{\pi^p}(t,x,a) - \vartheta^{\pi^p}(t,x,a')| &\leq |\vartheta^{\pi^p}(t,x,a) - v^{\pi^p}(t_p,x)| + |v^{\pi^p}(t_p,x) - \vartheta^{\pi^p}(t,x,a')| \\ &\leq 2C|\pi^p|^{\frac{1}{2}} \ . \end{aligned}$$

Sending p to ∞ we obtain that $\vartheta(t,x,a) = \vartheta(t,x,a')$. □

Corollary 4.1 *We have the identification: $\vartheta = v$, and the sequence $(v^\pi)_\pi$ also converges to v.*

Proof. We proceed in two steps.
Step 1. *The function ϑ is a supersolution to* (3.3). Since $\vartheta^{\pi_k}(T,.) = g$ for all $k \geq 1$, we first notice that $\vartheta(T,.) = g$. Next, since ϑ does not depend on the variable a, we have

$$\vartheta^\pi(t,x,a) \ \uparrow \ \vartheta(t,x) \quad \text{as} \ |\pi| \downarrow 0$$

for any $(t,x,a) \in [0,T] \times \mathbb{R}^d \times A$. Moreover, since the function ϑ is l.s.c, we have

$$\vartheta = \vartheta_* = \liminf_{|\pi|\to 0} {}_*\vartheta^\pi, \tag{4.3}$$

where

$$\liminf_{|\pi|\to 0} {}_*\vartheta^\pi(t,x,a) := \liminf_{\substack{|\pi| \to 0 \\ (t',x',a') \to (t,x,a) \\ t' < T}} \vartheta^\pi(t',x',a'), \qquad (t,x,a) \in [0,T] \times \mathbb{R}^d \times \mathbb{R}^q \ .$$

Fix now some $(t,x) \in [0,T] \times \mathbb{R}^d$ and $a \in A$ and $(p,q,M) \in \bar{J}^{2,-}\vartheta(t,x)$, the limiting parabolic subjet of ϑ at (t,x) (see definition in [9]). From standard stability results, there exists a sequence $(\pi_k, t_k, x_k, a_k, p_k, q_k, M_k)_k$ such that

$$(p_k, q_k, M_k) \ \in \ \bar{J}^{2,-}\vartheta^{\pi_k}(t_k, x_k, a_k)$$

for all $k \geq 1$ and

$$(t_k, x_k, a_k, \vartheta^{\pi_k}(t_k, x_k, a_k)) \ \longrightarrow \ (t, x, a, \vartheta(t,x,a)) \quad \text{as } k \to \infty, \ |\pi_k| \to 0.$$

From the viscosity supersolution property of ϑ^{π_k} to (3.9) in terms of subjets, we have

$$
\begin{aligned}
-p_k - b(x_k,a_k).q_k - \frac{1}{2}\mathrm{tr}(\sigma\sigma^\intercal(x_k,a_k)M_k) - f\big(x_k,a_k,\vartheta^{\pi_k}(t_k,x_k,a_k),\sigma^\intercal(x_k,a_k)q_k\big) & \\
- \int_A \big(\vartheta^{\pi_k}(t_k,x_k,a') - \vartheta^{\pi_k}(t_k,x_k,a_k)\big)\lambda(da') & \geq 0
\end{aligned}
$$

for all $k \geq 1$. Sending k to infinity and using (4.3), we get

$$
-p - b(x,a).q - \frac{1}{2}\mathrm{tr}(\sigma\sigma^\intercal(x,a)M) - f\big(x,a,\vartheta(t,x),\sigma^\intercal(x,a)q\big) \geq 0.
$$

Since a is arbitrary in A, this shows

$$
-p - \sup_{a\in A}\big[b(x,a).q + \frac{1}{2}\mathrm{tr}(\sigma\sigma^\intercal(x,a)M) + f\big(x,a,\vartheta(t,x),\sigma^\intercal(x,a)q\big)\big] \geq 0,
$$

i.e. the viscosity supersolution property of ϑ to (3.3).

Step 2. *Comparison.* Since the PDE (3.3) satisfies a comparison principle, we have from the previous step $\vartheta \geq v$, and we conclude with Lemma 4.1 that $\vartheta = v$. Finally, by definition (3.10) of v^π and from Lemma 4.1, we clearly have $\vartheta^\pi \leq v^\pi \leq v$, which also proves that $(v^\pi)_\pi$ converges to v. □

In terms of the discretely jump-constrained BSDE, the convergence result is formulated as follows:

Proposition 4.2 *We have $\mathcal{Y}_t^\pi \leq Y_t^\pi \leq Y_t$, $0 \leq t \leq T$, and*

$$
\mathbb{E}\Big[\sup_{t\in[0,T]}|Y_t - \mathcal{Y}_t^\pi|^2\Big] + \mathbb{E}\Big[\sup_{t\in[0,T]}|Y_t - Y_t^\pi|^2\Big] + \mathbb{E}\Big[\int_0^T |Z_t - \mathcal{Z}_t^\pi|^2 dt\Big] \rightarrow 0,
$$

as $|\pi|$ goes to zero.

Proof. Recall from (3.2) and (3.17) that we have the representation:

$$
Y_t = v(t,X_t), \quad Y_t^\pi = v^\pi(t,X_t,I_t), \quad \mathcal{Y}_t^\pi = \vartheta(t,X_t,I_t), \tag{4.4}
$$

and the first assertion of Lemma (4.1), we clearly have: $\mathcal{Y}_t^\pi \leq Y_t^\pi \leq Y_t$, $0 \leq t \leq T$. The convergence in $\mathcal{S}^2$ for $\mathcal{Y}^\pi$ to Y and Y^π to Y comes from the above representation (4.4), the pointwise convergence of ϑ^π and v^π to v in Corollary 4.1, and the dominated convergence theorem by recalling that $0 \leq (v - v^\pi)(t,x,a) \leq (v - \vartheta^\pi)(t,x,a) \leq v(t,x) \leq C(1+|x|)$. Let us now turn to the component Z. By definition (3.5)-(3.6)-(3.7) of the discretely jump-constrained BSDE we notice that $\mathcal{Y}^\pi$ can be written on $[0,T]$ as:

$$
\mathcal{Y}_t^\pi = g(X_T) + \int_t^T f(X_s,I_s,\mathcal{Y}_s^\pi,\mathcal{Z}_s^\pi) - \int_t^T \mathcal{Z}_s^\pi dW_s - \int_t^T\int_A \mathcal{U}_s^\pi(a)\tilde{\mu}(ds,da) + \mathcal{K}_T^\pi - \mathcal{K}_t^\pi,
$$

where $\mathcal{K}^\pi$ is the nondecreasing process defined by: $\mathcal{K}_t^\pi = \sum_{t_k\leq t}(Y_{t_k}^\pi - \mathcal{Y}_{t_k}^\pi)$, for $t \in [0,T]$. Denote by $\delta Y = Y - \mathcal{Y}^\pi$, $\delta Z = Z - \mathcal{Z}^\pi$, $\delta U = U - \mathcal{U}^\pi$ and $\delta K = K - \mathcal{K}^\pi$. From Itô's formula, Young Inequality and **(H2)**, there exists a constant C such that

$$\begin{aligned} & \mathbb{E}\Big[|\delta Y_t|^2\Big] + \frac{1}{2}\mathbb{E}\Big[\int_t^T |\delta Z_s|^2 ds\Big] + \frac{1}{2}\mathbb{E}\Big[\int_t^T |\delta U_s(a)|^2 \lambda(da) ds\Big] \\ \leq \;& C\int_t^T \mathbb{E}\Big[|\delta Y_s|^2\Big] ds + \frac{1}{\varepsilon}\mathbb{E}\Big[\sup_{s\in[0,T]} |\delta Y_s|^2\Big] + \varepsilon\mathbb{E}\Big[\big|\delta K_T - \delta K_t\big|^2\Big] \end{aligned} \tag{4.5}$$

for all $t \in [0,T]$, with ε a constant to be chosen later. From the definition of δK we have

$$\begin{aligned} \delta K_T - \delta K_t \;=\;& \delta Y_t - \int_t^T \big(f(X_s, I_s, Y_s, Z_s) - f(X_s, I_s, \mathcal{Y}_s^\pi, \mathcal{Z}_s^\pi)\big) ds \\ & + \int_0^T \delta Z_s dW_s + \int_t^T \int_A \delta U_s(a) \tilde{\mu}(ds, da)\;. \end{aligned}$$

Therefore, by (H2), we get the existence of a constant C' such that

$$\mathbb{E}\Big[\big|\delta K_T - \delta K_t\big|^2\Big] \;\leq\; C'\Big(\mathbb{E}\Big[\sup_{s\in[0,T]} |\delta Y_s|^2\Big] + \mathbb{E}\Big[\int_t^T |\delta Z_s|^2 ds\Big] + \mathbb{E}\Big[\int_t^T |\delta U_s(a)|^2 \lambda(da) ds\Big]\Big)$$

Taking $\varepsilon = \frac{C'}{4}$ and plugging this last inequality in (4.5), we get the existence of a constant C'' such that

$$\mathbb{E}\Big[\int_t^T |\delta Z_s|^2 ds\Big] + \mathbb{E}\Big[\int_t^T |\delta U_s(a)|^2 \lambda(da) ds\Big] \;\leq\; C''\Big(\mathbb{E}\Big[\sup_{s\in[0,T]} |\delta Y_s|^2\Big]\Big), \tag{4.6}$$

which shows the $L^2(W)$ convergence of $\mathcal{Z}^\pi$ to Z from the $\mathcal{S}^2$ convergence of $\mathcal{Y}^\pi$ to Y. □

4.2 Rate of convergence

We next provide an error estimate for the convergence of the discretely jump-constrained BSDE. We shall combine BSDE methods and PDE arguments adapted from the shaking coefficients approach of Krylov [17] and switching systems approximation of Barles, Jacobsen [4]. We make further assumptions:

(H1') The functions b and σ are uniformly bounded:

$$\sup_{x\in\mathbb{R}^d, a\in A} |b(x,a)| + |\sigma(x,a)| \;<\; \infty.$$

(H2') The function f does not depend on z: $f(x,a,y,z) = f(x,a,y)$ for all $(x,a,y,z) \in \mathbb{R}^d \times A \times \mathbb{R} \times \mathbb{R}^d$ and

(i) the functions $f(.,.,0)$ and g are uniformly bounded:

$$\sup_{x\in\mathbb{R}^d, a\in A} |f(x,a,0)| + |g(x)| \;<\; \infty,$$

(ii) for all $(x,a) \in \mathbb{R}^d \times A$, $y \mapsto f(x,a,y)$ is convex.

Under these assumptions, we obtain the rate of convergence for v^π and ϑ^π towards v.

Theorem 4.1 *Under* **(H1')** *and* **(H2')**, *there exists a constant C such that*

$$0 \quad \leq \quad v(t,x) - v^\pi(t,x,a) \quad \leq \quad v(t,x) - \vartheta^\pi(t,x,a) \quad \leq \quad C|\pi|^{\frac{1}{10}}$$

for all $(t,x,a) \in [0,T] \times \mathbb{R}^d \times A$ and all grid π with $|\pi| \leq 1$. Moreover, when $f(x,a)$ does not depend on y, the rate of convergence is improved to $|\pi|^{\frac{1}{6}}$.

Before proving this result, we give as corollary the rate of convergence for the discretely jump-constrained BSDE.

Corollary 4.2 *Under* **(H1')** *and* **(H2')**, *there exists a constant C such that*

$$\mathbb{E}\Big[\sup_{t\in[0,T]} |Y_t - \mathcal{Y}_t^\pi|^2\Big] + \mathbb{E}\Big[\sup_{t\in[0,T]} |Y_t - Y_t^\pi|^2\Big] + \mathbb{E}\Big[\int_0^T |Z_t - \mathcal{Z}_t^\pi|^2 dt\Big] \quad \leq \quad C|\pi|^{\frac{1}{5}}.$$

for all grid π with $|\pi| \leq 1$, and the above rate is improved to $|\pi|^{\frac{1}{3}}$ when $f(x,a)$ does not depend on y.

Proof. From the representation (4.4), and Theorem 4.1, we immediately have

$$\mathbb{E}\Big[\sup_{t\in[0,T]} |Y_t - \mathcal{Y}_t^\pi|^2\Big] + \mathbb{E}\Big[\sup_{t\in[0,T]} |Y_t - Y_t^\pi|^2\Big] \quad \leq \quad C|\pi|^{\frac{1}{5}}, \tag{4.7}$$

(resp. $|\pi|^{\frac{1}{3}}$ when $f(x,a)$ does not depend on y). Finally, the convergence rate for Z follows from the inequality (4.6). □

Remark 4.1 The above convergence rate $|\pi|^{\frac{1}{10}}$ is the optimal rate that one can prove in our generalized stochastic control context with fully nonlinear HJB equation by PDE approach and shaking coefficients technique, see [17], [4], [10] or [20]. However, this rate may not be the sharpest one. In the case of continuously reflected BSDEs, i.e. BSDEs with upper or lower constraint on Y, it is known that Y can be approximated by discretely reflected BSDEs, i.e. BSDEs where reflection on Y operates a finite set of times on the grid π, with a rate $|\pi|^{\frac{1}{2}}$ (see [1]). The standard arguments for proving this rate is based on the representation of the continuously (resp. discretely) reflected BSDE as optimal stopping problems where stopping is possible over the whole interval time (resp. only on the grid times). In our jump-constrained case, we know from [13] that the minimal solution to the BSDE with nonpositive jumps has the stochastic control representation (1.2) when $f(x,a)$ does not depend on y and z. We shall prove an analog representation for discretely jump-constrained BSDEs, and this helps to improve the rate of convergence from $|\pi|^{\frac{1}{10}}$ to $|\pi|^{\frac{1}{6}}$. □

The rest of this section is devoted to the proof of Theorem 4.1. We first consider the special case where $f(x,a)$ does not depend on y, and then address the case $f(x,a,y)$.

Proof of Theorem 4.1 in the case $f(x,a)$.
In the case where $f(x,a)$ does not depend on y,z, by (linear) Feynman-Kac formula for ϑ^π solution to (3.9), and by definition of v^π in (3.10), we have:

$$v^\pi(t_k,x) \quad = \quad \sup_{a\in A} \mathbb{E}\Big[\int_{t_k}^{t_{k+1}} f(X_t^{t_k,x,a}, I_t^{t_k,a})dt + v^\pi(t_{k+1}, X_{t_{k+1}}^{t_k,x,a})\Big], \quad k \leq n-1, \ x \in \mathbb{R}^d.$$

By induction, this dynamic programming relation leads to the following stochastic control problem with discrete time policies:

$$v^\pi(t_k, x) \;\; = \;\; \sup_{\alpha \in \mathcal{A}^\pi_{\mathbb{F}}} \mathbb{E}\Big[\int_{t_k}^T f(\bar{X}_t^{t_k,x,\alpha}, \bar{I}_t^\alpha) dt + g(\bar{X}_T^{t_k,x,\alpha}) \Big],$$

where $\mathcal{A}^\pi_{\mathbb{F}}$ is the set of discrete time processes $\alpha = (\alpha_{t_j})_{j \leq n-1}$, with α_{t_j} $\mathcal{F}_{t_j}$-measurable, valued in A, and

$$\begin{aligned} \bar{X}_t^{t_k,x,\alpha} &= x + \int_{t_k}^t b(\bar{X}_s^{t,x,\alpha}, \bar{I}_s^\alpha) ds + \int_{t_k}^t \sigma(\bar{X}_s^{t_k,x,\alpha}, \bar{I}_s^\alpha) dW_s, \quad t_k \leq t \leq T, \\ \bar{I}_t^\alpha &= \alpha_{t_j} + \int_{(t_j,t]} \int_A (a - \bar{I}_{s^-}^\alpha) \mu(ds, da), \quad t_j \leq t < t_{j+1}, \; j \leq n-1. \end{aligned}$$

In other words, $v^\pi(t_k, x)$ corresponds to the value function for a stochastic control problem where the controller can act only at the dates t_j of the grid π, and then let the regime of the coefficients of the diffusion evolve according to the Poisson random measure μ. Let us introduce the following stochastic control problem with piece-wise constant control policies:

$$\tilde{v}^\pi(t_k, x) \;\; = \;\; \sup_{\alpha \in \mathcal{A}^\pi_{\mathbb{F}}} \mathbb{E}\Big[\int_{t_k}^T f(\tilde{X}_t^{t_k,x,\alpha}, \tilde{I}_t^\alpha) dt + g(\tilde{X}_T^{t_k,x,\alpha}) \Big],$$

where for $\alpha = (\alpha_{t_j})_{j \leq n-1} \in \mathcal{A}^\pi_{\mathbb{F}}$:

$$\begin{aligned} \tilde{X}_t^{t_k,x,\alpha} &= x + \int_{t_k}^t b(\tilde{X}_s^{t,x,\alpha}, \tilde{I}_s^\alpha) ds + \int_{t_k}^t \sigma(\tilde{X}_s^{t_k,x,\alpha}, \tilde{I}_s^\alpha) dW_s, \quad t_k \leq t \leq T, \\ \tilde{I}_t^\alpha &= \alpha_{t_j}, \qquad t_j \leq t < t_{j+1}, \; j \leq n-1. \end{aligned}$$

It is shown in [16] that $\tilde{v}^\pi$ approximates the value function v for the controlled diffusion problem (1.2), solution to the HJB equation (3.3), with a rate $|\pi|^{\frac{1}{6}}$:

$$0 \;\leq\; v(t_k, x) - \tilde{v}^\pi(t_k, x) \;\; \leq \;\; C|\pi|^{\frac{1}{6}}, \tag{4.8}$$

for all $t_k \in \pi$, $x \in \mathbb{R}^d$. Now, recalling that A is compact and $\lambda(A) < \infty$, it is clear that there exists some positive constant C such that for all $\alpha \in \mathcal{A}^\pi_{\mathbb{F}}$, $j \leq n-1$:

$$\mathbb{E}\Big[\sup_{t \in [t_j, t_{j+1})} |\bar{I}_t^\alpha - \tilde{I}_t^\alpha|^2 \Big] \;\; \leq \;\; C|\pi|,$$

and then by standard arguments under Lipschitz condition on b, σ:

$$\mathbb{E}\Big[\sup_{t \in [t_j, t_{j+1}]} |\bar{X}_t^{t_k,x,\alpha} - \tilde{X}_t^{t_k,x,\alpha}|^2 \Big] \;\; \leq \;\; C|\pi|, \quad k \leq j \leq n-1, \; x \in \mathbb{R}^d.$$

By the Lipschitz conditions on f and g, it follows that

$$|v^\pi(t_k, x) - \tilde{v}^\pi(t_k, x)| \;\; \leq \;\; C|\pi|^{\frac{1}{2}},$$

and thus with (4.8):

$$0 \;\leq\; \sup_{x \in \mathbb{R}^d} (v - v^\pi)(t_k, x) \;\; \leq \;\; C|\pi|^{\frac{1}{6}}.$$

Finally, by combining with the estimate in Lemma 4.2, which gives actually under **(H2')**(i):

$$|\vartheta^\pi(t,x,a) - v^\pi(t_{k+1},x)| \;\; \leq \;\; C|\pi|^{\frac{1}{2}}, \quad t \in [t_k,t_{k+1}), (x,a) \in \mathbb{R}^d \times A,$$

together with the 1/2-Hölder property of v in time (see (3.4)), we obtain:

$$\sup_{(t,x,a)\in[0,T]\times\mathbb{R}^d\times A} (v-\vartheta^\pi)(t,x,a) \;\; \leq \;\; C(|\pi|^{\frac{1}{6}} + |\pi|^{\frac{1}{2}}) \;\leq\; C|\pi|^{\frac{1}{6}},$$

for $|\pi| \leq 1$. This ends the proof. $\square$

Let us now turn to the case where $f(x,a,y)$ may also depend on y. We cannot rely anymore on the convergence rate result in [16]. Instead, recalling that A is compact and since σ, b and f are Lipschitz in (x,a), we shall apply the switching system method of Barles and Jacobsen [4], which is a variation of the shaken coefficients method and smoothing technique used in Krylov [17], in order to obtain approximate smooth subsolution to (3.3). By Lemmas 3.3 and 3.4 in [4], one can find a family of smooth functions $(w_\varepsilon)_{0<\varepsilon\leq 1}$ on $[0,T]\times\mathbb{R}^d$ such that:

$$\sup_{[0,T]\times\mathbb{R}^d} |w_\varepsilon| \;\; \leq \;\; C, \tag{4.9}$$

$$\sup_{[0,T]\times\mathbb{R}^d} |w_\varepsilon - w| \;\; \leq \;\; C\varepsilon^{\frac{1}{3}}, \tag{4.10}$$

$$\sup_{[0,T]\times\mathbb{R}^d} |\partial_t^{\beta_0} D^\beta w_\varepsilon| \;\; \leq \;\; C\varepsilon^{1-2\beta_0-\sum_{i=1}^d \beta^i}, \;\; \beta_0 \in \mathbb{N}, \; \beta = (\beta^1,\ldots,\beta^d) \in \mathbb{N}^d, \tag{4.11}$$

for some positive constant C independent of ε, and by convexity of f in **(H2')**(ii), for any $\varepsilon \in (0,1]$, $(t,x) \in [0,T]\times\mathbb{R}^d$, there exists $a_{t,x,\varepsilon} \in A$ such that:

$$-\mathcal{L}^{a_{t,x,\varepsilon}} w_\varepsilon(t,x) - f(x,a_{t,x,\varepsilon},w_\varepsilon(t,x)) \;\; \geq \;\; 0. \tag{4.12}$$

Recalling the definition of the operator $\mathbb{T}_\pi^k$ in (3.11), we define for any function φ on $[0,T]\times\mathbb{R}^d\times A$, Lipschitz in (x,a):

$$\mathbb{T}_\pi[\varphi](t,x,a) \;\; := \;\; \mathbb{T}_\pi^k[\varphi(t_{k+1},.,.)](t,x,a), \quad t \in [t_k,t_{k+1}), \; (x,a) \in \mathbb{R}^d \times A,$$

for $k = 0,\ldots,n-1$, and

$$\begin{aligned}\mathbb{S}_\pi[\varphi](t,x,a) \;\; := \;\; & \frac{1}{|\pi|}\Big[\varphi(t,x) - \mathbb{T}_\pi[\varphi](t,x,a) \\ & \qquad +(t_{k+1}-t)\big(\mathcal{L}^a\varphi(t,x) + f(x,a,\varphi(t,x))\big)\Big],\end{aligned}$$

for $(t,x,a) \in [t_k,t_{k+1})\times\mathbb{R}^d\times A$, $k \leq n-1$.

We have the following key error bound on $\mathbb{S}_\pi$.

Lemma 4.3 *Let* **(H1')** *and* **(H2')***(i) hold. There exists a constant C such that*

$$|\mathbb{S}_\pi[\varphi_\varepsilon](t,x,a)| \;\; \leq \;\; C\Big(|\pi|^{\frac{1}{2}}(1+\varepsilon^{-1}) + |\pi|\varepsilon^{-3}\Big), \quad (t,x,a) \in [0,T]\times\mathbb{R}^d\times A,$$

for any family $(\varphi_\varepsilon)_\varepsilon$ of smooth functions on $[0,T]\times\mathbb{R}^d$ satisfying (4.9) *and* (4.11).

Proof. Fix $(t, x, a) \in [0, T] \times \mathbb{R}^d \times A$. If $t = T$, we have $|\mathbb{S}_\pi[\varphi_\varepsilon](t, x, a)| = 0$. Suppose that $t < T$ and fix $k \le n$ such that $t \in [t_k, t_{k+1})$. Given a smooth function φ_ε satisfying (4.9) and (4.11), we split:

$$|\mathbb{S}_\pi[\varphi_\varepsilon](t, x, a)| \quad \le \quad A_\varepsilon(t, x, a) + B_\varepsilon(t, x, a),$$

where

$$A_\varepsilon(t, x, a) \quad := \quad \frac{1}{|\pi|}\Big|\mathbb{T}_\pi[\varphi_\varepsilon](t, x, a) - \mathbb{E}\big[\varphi_\varepsilon(t_{k+1}, X^{t,x,a}_{t_{k+1}})\big] - (t_{k+1} - t) f(x, a, \varphi_\varepsilon(t, x))\Big| \,,$$

and

$$B_\varepsilon(t, x, a) \quad := \quad \frac{1}{|\pi|}\Big|\mathbb{E}\big[\varphi_\varepsilon(t_{k+1}, X^{t,x,a}_{t_{k+1}})\big] - \varphi_\varepsilon(t, x) - (t_{k+1} - t)\mathcal{L}^a \varphi_\varepsilon(t, x)\Big|,$$

and we study each term A_ε and B_ε separately.

1. *Estimate on $A_\varepsilon(t, x, a)$.*
Define $(Y^{\varphi_\varepsilon}, Z^{\varphi_\varepsilon}, U^{\varphi_\varepsilon})$ as the solution to the BSDE on $[t, t_{k+1}]$:

$$\begin{aligned} Y^{\varphi_\varepsilon}_s \quad &= \quad \varphi_\varepsilon(t_{k+1}, X^{t,x,a}_{t_{k+1}}) + \int_s^{t_{k+1}} f(X^{t,x,a}_r, I^{t,a}_r, Y^{\varphi_\varepsilon}_r) dr \\ &\quad - \int_s^{t_{k+1}} Z^{\varphi_\varepsilon}_r dW_r - \int_s^{t_{k+1}} \int_A U^{\varphi_\varepsilon}_r(a) \tilde{\mu}(dr, da) \,, \quad s \in [t, t_{k+1}]. \end{aligned} \tag{4.13}$$

From Theorems 3.4 and 3.5 in [3], we have $Y^{\varphi_\varepsilon}_t = \mathbb{T}_\pi[\varphi_\varepsilon](t, x, a)$, and by taking expectation in (4.13), we thus get:

$$Y^{\varphi_\varepsilon}_t \; = \; \mathbb{T}_\pi[\varphi_\varepsilon](t, x, a) \quad = \quad \mathbb{E}\big[\varphi_\varepsilon(t_{k+1}, X^{t,x,a}_{t_{k+1}})\big] + \mathbb{E}\Big[\int_t^{t_{k+1}} f(X^{t,x,a}_s, I^{t,a}_s, Y^{\varphi_\varepsilon}_s) ds\Big]$$

and so:

$$\begin{aligned} A_\varepsilon(t, x, a) \quad &\le \quad \frac{1}{|\pi|}\mathbb{E}\Big[\int_t^{t_{k+1}} \big|f(X^{t,x,a}_s, I^{t,a}_s, Y^{\varphi_\varepsilon}_s) - f(x, a, \varphi_\varepsilon(t, x))\big| ds\Big] \\ &\le \quad C\Big(\mathbb{E}\big[\sup_{s \in [t, t_{k+1}]} |X^{t,x,a}_s - x| + |I^{t,a}_s - a|\big] + \mathbb{E}\big[\sup_{s \in [t, t_{k+1}]} |Y^{\varphi_\varepsilon}_s - \varphi_\varepsilon(t, x)|\big]\Big), \end{aligned}$$

by the Lipschitz continuity of f. From standard estimate for SDE, we have (recall that the coefficients b and σ are bounded under **(H1')** and A is compact):

$$\mathbb{E}\big[\sup_{s \in [t, t_{k+1}]} |X^{t,x,a}_s - x| + |I^{t,a}_s - a|\big] \quad \le \quad C|\pi|^{\frac{1}{2}}. \tag{4.14}$$

Moreover, by (4.13), the boundedness condition in **(H2')**(i) together with the Lipschitz condition of f, and Burkholder-Davis-Gundy inequality, we have:

$$\begin{aligned} \mathbb{E}\Big[\sup_{s \in [t, t_{k+1}]} |Y^{\varphi_\varepsilon}_s - \varphi_\varepsilon(t, x)|\Big] \quad &\le \quad \mathbb{E}\big[|\varphi_\varepsilon(t_{k+1}, X^{t,x,a}_{t_{k+1}}) - \varphi_\varepsilon(t, x)|\big] \\ &\quad + C|\pi|\mathbb{E}\big[1 + \sup_{s \in [t, t_{k+1}]} |Y^{\varphi_\varepsilon}_s|\big] \\ &\quad + C|\pi|\Big(\mathbb{E}\big[\sup_{s \in [t, t_{k+1}]} |Z^{\varphi_\varepsilon}_s|^2\big] + \mathbb{E}\big[\sup_{s \in [t, t_{k+1}]} \int_A |U^{\varphi_\varepsilon}_s(a)|^2 \lambda(da)\big]\Big). \end{aligned}$$

From standard estimate for the BSDE (4.13), we have:

$$\mathbb{E}\big[\sup_{s\in[t,t_{k+1}]}|Y_s^{\varphi_\varepsilon}|^2\big] \quad \leq \quad C,$$

for some positive constant C depending only on the Lipschitz constant of f, the upper bound of $|f(x,a,0,0)|$ in **(H2')**(i), and the upper bound of $|\varphi_\varepsilon|$ in (4.9). Moreover, from the estimate in Proposition 4.2 in [7] about the coefficients Z^{φ_ε} and U^{φ_ε} of the BSDE with jumps (4.13), there exists some constant C depending only on the Lipschitz constant of b,σ,f, and of the Lipschitz constant of $\varphi_\varepsilon(t_{k+1},.)$ (which does not depend on ε by (4.11)), such that:

$$\mathbb{E}\big[\sup_{s\in[t,t_{k+1}]}|Z_s^{\varphi_\varepsilon}|^2\big]+\mathbb{E}\big[\sup_{s\in[t,t_{k+1}]}\int_A|U_s^{\varphi_\varepsilon}(a)|^2\lambda(da)\big] \quad \leq \quad C.$$

From (4.11), we then have:

$$\begin{aligned}\mathbb{E}\Big[\sup_{s\in[t,t_{k+1}]}|Y_s^{\varphi_\varepsilon}-\varphi_\varepsilon(t,x)|\Big] &\leq C\big(|t_{k+1}-t|\varepsilon^{-1}+\mathbb{E}\big[|X_{t_{k+1}}^{t,x,a}-x|\big]+|\pi|\big)\\ &\leq C|\pi|^{\frac{1}{2}}\big(1+\varepsilon^{-1}\big),\end{aligned}$$

by (4.14). This leads to the error bound for $A_\varepsilon(t,x,a)$:

$$A_\varepsilon(t,x,a) \quad \leq \quad C|\pi|^{\frac{1}{2}}(1+\varepsilon^{-1}).$$

2. *Estimate on $B_\varepsilon(t,x,a)$.*
From Itô's formula we have

$$\begin{aligned}B_\varepsilon(t,x,a) &= \frac{1}{|\pi|}\Big|\mathbb{E}\Big[\int_t^{t_{k+1}}\big(\mathcal{L}^{I_s^{t,a}}\varphi_\varepsilon(s,X_s^{t,x,a})-\mathcal{L}^a\varphi_\varepsilon(t,x)\big)ds\Big]\Big|\\ &\leq B_\varepsilon^1(t,x,a)+B_\varepsilon^2(t,x,a)\end{aligned}$$

where

$$\begin{aligned}B_\varepsilon^1(t,x,a) &= \frac{1}{|\pi|}\mathbb{E}\Big[\int_t^{t_{k+1}}\Big|\big(b(X_s^{t,x,a},I_s^{t,a})-b(x,a)\big).D_x\varphi_\varepsilon(s,X_s^{t,x,a})\\ &\quad+\frac{1}{2}\mathrm{tr}\big[\big(\sigma\sigma^\intercal(X_s^{t,x,a},I_s^{t,a})-\sigma\sigma^\intercal(x,a)\big)D_x^2\varphi_\varepsilon(t,x)\big]\Big|ds\Big]\end{aligned}$$

and

$$B_\varepsilon^2(t,x,a) \quad = \quad \frac{1}{|\pi|}\mathbb{E}\Big[\int_t^{t_{k+1}}\big|\tilde{\mathcal{L}}_{t,x}^a\varphi_\varepsilon(s,X_s^{t,x,a})-\tilde{\mathcal{L}}_{t,x}^a\varphi_\varepsilon(t,x)\big|ds\Big],$$

with $\tilde{\mathcal{L}}_{t,x}^a$ defined by

$$\tilde{\mathcal{L}}_{t,x}^a\varphi_\varepsilon(t',x') \quad = \quad \frac{\partial\varphi_\varepsilon}{\partial t}(t',x')+b(x,a).D_x\varphi_\varepsilon(t',x')+\frac{1}{2}\mathrm{tr}\big(\sigma\sigma^\intercal(x,a)D_x^2\varphi_\varepsilon(t',x')\big).$$

Under **(H1)**, **(H1')**, and by (4.11), we have

$$\begin{aligned}B_\varepsilon^1(t,x,a) &\leq C(1+\varepsilon^{-1})\mathbb{E}\Big[\sup_{s\in[t,t_{k+1}]}|X_s^{t,x,a}-x|+|I_s^{t,a}-a|\Big]\\ &\leq C(1+\varepsilon^{-1})|\pi|^{\frac{1}{2}},\end{aligned}$$

where we used again (4.14). On the other hand, since φ_ε is smooth, we have from Itô's formula

$$B^2_\varepsilon(t,x,a) \;=\; \frac{1}{|\pi|}\mathbb{E}\Big[\int_t^{t_{k+1}}\Big|\int_t^s \mathcal{L}^{I_r^{t,a}}\tilde{\mathcal{L}}^a_{t,x}\phi(r,X_r^{t,x,a})dr\Big|ds\Big] .$$

Under **(H1')** and by (4.11), we then see that

$$B^2_\varepsilon(t,x,a) \;\leq\; C|\pi|\varepsilon^{-3},$$

and so:

$$B_\varepsilon(t,x,a) \;\leq\; C\Big(|\pi|^{\frac{1}{2}}(1+\varepsilon^{-1})+|\pi|\varepsilon^{-3}\Big).$$

Together with the estimate for $A_\varepsilon(t,x,a)$, this proves the error bound for $|\mathbb{S}_\pi[\varphi_\varepsilon](t,x,a)|$. □

We next state a maximum principle type result for the operator $\mathbb{T}_\pi$.

Lemma 4.4 *Let φ and ψ be two functions on $[0,T]\times\mathbb{R}^d\times A$, Lipschitz in (x,a). Then, there exists some positive constant C independent of π such that*

$$\sup_{(t,x,a)\in[t_k,t_{k+1}]\times\mathbb{R}^d\times A}(\mathbb{T}_\pi[\varphi]-\mathbb{T}_\pi[\psi])(t,x,a) \;\leq\; e^{C|\pi|}\sup_{(x,a)\in\mathbb{R}^d\times A}(\varphi-\psi)(t_{k+1},x,a) ,$$

for all $k=0,\ldots,n-1$.

Proof. Fix $k\leq n-1$, and set

$$M \;:=\; \sup_{(x,a)\in\mathbb{R}^d\times A}(\varphi-\psi)(t_{k+1},x,a).$$

We can assume w.l.o.g. that $M<\infty$ since otherwise the required inequality is trivial. Let us denote by Δv the function

$$\Delta v(t,x,a) \;=\; \mathbb{T}_\pi[\varphi](t,x,a)-\mathbb{T}_\pi[\psi](t,x,a),$$

for all $(t,x,a)\in[t_k,t_{k+1}]\times\mathbb{R}^d\times A$. By definition of $\mathbb{T}_\pi$, and from the Lipschitz condition of f, we see that Δv is a viscosity subsolution to

$$\begin{cases} -\mathcal{L}^a\Delta v(t,x,a)-C\big(|\Delta v(t,x,a)|+|D\Delta v(t,x,a)|\big) \\ -\int_A\big(\Delta v(t,x,a')-\Delta v(t,x,a)\big)\lambda(da')=0, \quad \text{for } (t,x,a)\in[t_k,t_{k+1})\times\mathbb{R}^d\times A, \\ \Delta v(t_{k+1},x,a)\leq M , \qquad \text{for } (x,a)\in\mathbb{R}^d\times A . \end{cases} \tag{4.15}$$

Then, we easily check that the function Φ defined by

$$\Phi(t,x,a) \;=\; Me^{C(t_{k+1}-t)} , \quad (t,x,a)\in[t_k,t_{k+1}]\times\mathbb{R}^d\times A ,$$

is a solution to

$$\begin{cases} -\mathcal{L}^a\Phi(t,x,a)-C\big(|\Phi(t,x,a)|+|D\Phi(t,x,a)|\big) \\ -\int_A\big(\Phi(t,x,a')-\Phi(t,x,a)\big)\lambda(da') \;=\; 0, \quad \text{for } (t,x,a)\in[t_k,t_{k+1})\times\mathbb{R}^d\times A, \\ \Phi(t_{k+1},x,a) \;=\; M , \qquad \text{for } (x,a)\in\mathbb{R}^d\times A . \end{cases} \tag{4.16}$$

From the comparison theorem in [3] for viscosity solutions of semi-linear IPDEs, we get that $\Delta v \leq \Phi$ on $[t_k, t_{k+1}] \times \mathbb{R}^d \times A$, which proves the required inequality. $\square$

Proof of Theorem 4.1. By (3.10) and (3.16), we observe that v^π is a fixed point of $\mathbb{T}_\pi$, i.e.

$$\mathbb{T}_\pi[v^\pi] \;=\; v^\pi.$$

On the other hand, by (4.12), and the estimate of Lemma 4.3 applied to w_ε, we have:

$$w_\varepsilon(t,x) - \mathbb{T}_\pi[w_\varepsilon](t,x,a_{t,x,\varepsilon}) \;\leq\; |\pi|\mathbb{S}_\pi[w_\varepsilon](t,x,a_{t,x,\varepsilon}) \;\leq\; C|\pi|\bar{S}(\pi,\varepsilon)$$

where we set: $\bar{S}(\pi,\varepsilon) = (|\pi|^{\frac{3}{2}}(1+\varepsilon^{-1}) + |\pi|^2\varepsilon^{-3})$. Fix $k \leq n-1$. By Lemma 4.4, we then have for all $t \in [t_k, t_{k+1}]$, $x \in \mathbb{R}^d$:

$$\begin{aligned} w_\varepsilon(t,x) - v^\pi(t,x,a_{t,x,\varepsilon}) &= w_\varepsilon(t,x) - \mathbb{T}_\pi[w_\varepsilon](t,x,a_{t,x,\varepsilon}) + (\mathbb{T}_\pi[w_\varepsilon] - \mathbb{T}_\pi[v^\pi])(t,x,a_{t,x,\varepsilon}) \\ &\leq C|\pi|\bar{S}(\pi,\varepsilon) + e^{C|\pi|} \sup_{(x,a)\in\mathbb{R}^d\times A} (w_\varepsilon - v^\pi)(t_{k+1},x,a). \end{aligned} \tag{4.17}$$

Recalling by its very definition that v^π does not depend on $a \in A$ on the grid times of π, and denoting then $M_k := \sup_{x\in\mathbb{R}^d}(w_\varepsilon - v^\pi)(t_k,x)$, we have by (4.17) the relation:

$$M_k \;\leq\; C|\pi|\bar{S}(\pi,\varepsilon) + e^{C|\pi|}M_{k+1}.$$

By induction, this yields:

$$\begin{aligned} \sup_{x\in\mathbb{R}^d}(w_\varepsilon - v^\pi)(t_k,x) &\leq C\frac{e^{Cn|\pi|}-1}{e^{C|\pi|}-1}|\pi|\bar{S}(\pi,\varepsilon) + e^{Cn|\pi|}\sup_{x\in\mathbb{R}^d}(w_\varepsilon - v^\pi)(T,x) \\ &\leq C\bar{S}(\pi,\varepsilon) + C\sup_{x\in\mathbb{R}^d}(w_\varepsilon - v)(T,x), \end{aligned}$$

since $n|\pi|$ is bounded and $v(T,x) = v^\pi(T,x)$ $(= g(x))$. From (4.10), we then get:

$$\sup_{x\in\mathbb{R}^d}(v - v^\pi)(t_k,x) \;\leq\; C\big(\varepsilon^{\frac{1}{3}} + |\pi|^{\frac{1}{2}}(1+\varepsilon^{-1}) + |\pi|\varepsilon^{-3}\big).$$

By minimizing the r.h.s of this estimate with respect to ε, this leads to the error bound when taking $\varepsilon = |\pi|^{\frac{3}{10}} \leq 1$:

$$\sup_{x\in\mathbb{R}^d}(v - v^\pi)(t_k,x) \;\leq\; C|\pi|^{\frac{1}{10}}.$$

Finally, by combining with the estimate in Lemma 4.2, which gives actually under **(H2')**(i):

$$|\vartheta^\pi(t,x,a) - v^\pi(t_{k+1},x)| \;\leq\; C|\pi|^{\frac{1}{2}}, \quad t \in [t_k, t_{k+1}), (x,a) \in \mathbb{R}^d \times A,$$

together with the 1/2-Hölder property of v in time (see (3.4)), we obtain:

$$\sup_{(t,x,a)\in[0,T]\times\mathbb{R}^d\times A}(v - \vartheta^\pi)(t,x,a) \;\leq\; C(|\pi|^{\frac{1}{10}} + |\pi|^{\frac{1}{2}}) \;\leq\; C|\pi|^{\frac{1}{10}}.$$

This ends the proof. $\square$

5 Approximation scheme for jump-constrained BSDE and stochastic control problem

We consider the discrete time approximation of the discretely jump-constrained BSDE in the case where $f(x,a,y)$ does not depend on z, and define the scheme $(\bar Y^\pi, \bar{\mathcal{Y}}^\pi, \bar{\mathcal{Z}}^\pi)$ by induction on the grid $\pi = \{t_0 = 0 < \ldots < t_k < \ldots < t_n = T\}$ by:

$$\begin{cases} \bar Y^\pi_T \;=\; \bar{\mathcal{Y}}^\pi_T \;=\; g(\bar X^\pi_T) \\ \qquad\quad \bar{\mathcal{Y}}^\pi_{t_k} \;=\; \mathbb{E}_{t_k}\big[\bar Y^\pi_{t_{k+1}}\big] + f(\bar X^\pi_{t_k}, I_{t_k}, \bar{\mathcal{Y}}^\pi_{t_k})\Delta t_k \\ \qquad\quad \bar Y^\pi_{t_k} \;=\; \operatorname*{ess\,sup}_{a\in A} \mathbb{E}_{t_k,a}\big[\bar{\mathcal{Y}}^\pi_{t_k}\big], \quad k = 0,\ldots,n-1, \end{cases} \tag{5.1}$$

where $\Delta t_k = t_{k+1} - t_k$, $\mathbb{E}_{t_k}[.]$ stands for $\mathbb{E}[.|\mathcal{F}_{t_k}]$, and $\mathbb{E}_{t_k,a}[.]$ for $\mathbb{E}[.|\mathcal{F}_{t_k}, I_{t_k} = a]$.

By induction argument, we easily see that $\bar{\mathcal{Y}}^\pi_{t_k}$ is a deterministic function of $(\bar X^\pi_{t_k}, I_{t_k})$, while $\bar Y^\pi_{t_k}$ is a deterministic function of $\bar X^\pi_{t_k}$, for $k = 0,\ldots,n$, and by the Markov property of the process $(\bar X^\pi, I)$, the conditional expectations in (5.1) can be replaced by the corresponding regressions:

$$\mathbb{E}_{t_k}\big[\bar Y^\pi_{t_{k+1}}\big] \;=\; \mathbb{E}\big[\bar Y^\pi_{t_{k+1}}\big|\bar X^\pi_{t_k}, I_{t_k}\big] \quad \text{and} \quad \mathbb{E}_{t_k,a}\big[\bar{\mathcal{Y}}^\pi_{t_k}\big] \;=\; \mathbb{E}\big[\bar{\mathcal{Y}}^\pi_{t_k}\big|\bar X^\pi_{t_k}, I_{t_k} = a\big].$$

We then have:

$$\bar{\mathcal{Y}}^\pi_{t_k} \;=\; \bar\vartheta^\pi_k(\bar X^\pi_{t_k}, I_{t_k}), \qquad Y^\pi_{t_k} \;=\; \bar v^\pi_k(\bar X^\pi_{t_k}),$$

for some sequence of functions $(\bar\vartheta^\pi_k)_k$ and $(\bar v^\pi_k)_k$ defined respectively on $\mathbb{R}^d \times A$ and $\mathbb{R}^d$ by backward induction:

$$\begin{cases} \bar v^\pi_n(x,a) \;=\; \bar\vartheta^\pi_n(x) \;=\; g(x) \\ \qquad\quad \bar\vartheta^\pi_k(x,a) \;=\; \mathbb{E}\Big[\bar v^\pi_{k+1}(\bar X^\pi_{t_{k+1}}, I_{t_{k+1}})\big|(\bar X^\pi_{t_k}, I_{t_k}) = (x,a)\Big] + f(x,a,\bar\vartheta^\pi_k(x,a))\Delta t_k \\ \qquad\quad \bar v^\pi_k(x) \;=\; \sup_{a\in A} \bar\vartheta^\pi_k(x,a), \quad k = 0,\ldots,n-1. \end{cases} \tag{5.2}$$

There are well-known different methods (Longstaff-Schwartz least square regression, quantization, Malliavin integration by parts, see e.g. [1], [12], [8]) for computing the above conditional expectations, and so the functions $\bar\vartheta^\pi_k$ and $\bar v^\pi_k$. It appears that in our context, the simulation-regression method on basis functions defined on $\mathbb{R}^d \times A$, is quite suitable since it allows us to derive at each time step $k \leq n-1$, a functional form $\hat a_k(x)$, which attains the supremum over A in $\bar\vartheta^\pi_k(x,a)$. We shall see later in this section that the feedback control $(\hat a_k(x))_k$ provides an approximation of the optimal control for the HJB equation associated to a stochastic control problem when $f(x,a)$ does not depend on y. We refer to our companion paper [14] for the details about the computation of functions $\bar\vartheta^\pi_k$, $\bar v^\pi_k$, $\hat a_k$ by simulation-regression methods, and the associated error analysis.

5.1 Error estimate for the discrete time scheme

The main result of this section is to state an error bound between the component Y^π of the discretely jump-constrained BSDE and the solution $(\bar Y^\pi, \bar{\mathcal{Y}}^\pi)$ to the above discrete time scheme.

Theorem 5.1 *There exists some constant C such that:*

$$\mathbb{E}\Big[\big|Y_{t_k}^\pi - \bar{Y}_{t_k}^\pi\big|^2\Big] + \sup_{t\in(t_k,t_{k+1}]}\mathbb{E}\Big[\big|Y_t^\pi - \bar{Y}_{t_{k+1}}^\pi\big|^2\Big] + \sup_{t\in[t_k,t_{k+1})}\mathbb{E}\Big[\big|Y_t^\pi - \bar{\mathcal{Y}}_{t_k}^\pi\big|^2\Big] \;\;\leq\;\; C|\pi|,$$

for all $k = 0,\dots,n-1$.

The above convergence rate $|\pi|^{\frac{1}{2}}$ in the L^2-norm for the discretization of the discretely jump-constrained BSDE is the same as for standard BSDE, see [8], [21]. By combining with the convergence result in Section 4, we finally obtain an estimate on the error due to the discrete time approximation of the minimal solution Y to the BSDE with nonpositive jumps. We split the error between the positive and negative parts:

$$\begin{aligned}
\mathrm{Err}_+^\pi(Y) &:= \max_{k\leq n-1}\Big(\mathbb{E}\Big[\big(Y_{t_k} - \bar{Y}_{t_k}^\pi\big)_+^2\Big] + \sup_{t\in(t_k,t_{k+1}]}\mathbb{E}\Big[\big(Y_t - \bar{Y}_{t_{k+1}}^\pi\big)_+^2\Big] + \sup_{t\in[t_k,t_{k+1})}\mathbb{E}\Big[\big(Y_t - \bar{\mathcal{Y}}_{t_k}^\pi\big)_+^2\Big]\Big)^{\frac{1}{2}}\\
\mathrm{Err}_-^\pi(Y) &:= \max_{k\leq n-1}\Big(\mathbb{E}\Big[\big(Y_{t_k} - \bar{Y}_{t_k}^\pi\big)_-^2\Big] + \sup_{t\in(t_k,t_{k+1}]}\mathbb{E}\Big[\big(Y_t - \bar{Y}_{t_{k+1}}^\pi\big)_-^2\Big] + \sup_{t\in[t_k,t_{k+1})}\mathbb{E}\Big[\big(Y_t - \bar{\mathcal{Y}}_{t_k}^\pi\big)_-^2\Big]\Big)^{\frac{1}{2}}.
\end{aligned}$$

Corollary 5.1 *We have:*

$$\mathrm{Err}_-^\pi(Y) \;\;\leq\;\; C|\pi|^{\frac{1}{2}}.$$

Moreover, under **(H1')** *and* **(H2')**,

$$\mathrm{Err}_+^\pi(Y) \;\;\leq\;\; C|\pi|^{\frac{1}{10}},$$

and when $f(x,a)$ does not depend on y:

$$\mathrm{Err}_+^\pi(Y) \;\;\leq\;\; C|\pi|^{\frac{1}{6}}.$$

Proof. Recall from Proposition 4.2 that $\mathcal{Y}_t^\pi \leq Y_t^\pi \leq Y_t$, $0 \leq t \leq T$. Then, we have: $(Y_{t_k} - \bar{Y}_{t_k}^\pi)_- \leq |Y_{t_k}^\pi - \bar{Y}_{t_k}^\pi|$, $(Y_t - \bar{Y}_{t_{k+1}}^\pi)_- \leq |Y_t^\pi - \bar{Y}_{t_{k+1}}^\pi|$, and $(Y_{t_k} - \bar{\mathcal{Y}}_{t_k}^\pi)_- \leq |Y_{t_k}^\pi - \bar{\mathcal{Y}}_{t_k}^\pi|$, for all $k \leq n-1$, and $t \in [0,T]$. The error bound on $\mathrm{Err}_-^\pi(Y)$ follows then from the estimation in Theorem 5.1. The error bound on $\mathrm{Err}_-^\pi(Y)$ follows from Corollary 4.2 and Theorem 5.1. □

Remark 5.1 In the particular case where f depends only on (x,a), our discrete time approximation scheme is a probabilistic scheme for the fully nonlinear HJB equation associated to the stochastic control problem (1.2). As in [17], [4] or [10], we have non symmetric bounds on the rate of convergence. For instance, in [10], the authors obtained a convergence rate $|\pi|^{\frac{1}{4}}$ on one side and $|\pi|^{\frac{1}{10}}$ on the other side, while we improve the rate to $|\pi|^{\frac{1}{2}}$ for one side, and $|\pi|^{\frac{1}{6}}$ on the other side. This induces a global error $\mathrm{Err}^\pi(Y) = \mathrm{Err}_+^\pi(Y)$ + $\mathrm{Err}_-^\pi(Y)$ of order $|\pi|^{\frac{1}{6}}$, which is derived without any non degeneracy condition on the controlled diffusion coefficient. □

Proof of Theorem 5.1.

Let us introduce the continuous time version of (5.1). By the martingale representation theorem, there exists $\tilde{\mathcal{Z}}^\pi \in L^2(W)$ and $\tilde{\mathcal{U}}^\pi \in L^2(\tilde{\mu})$ such that

$$\bar{Y}^\pi_{t_{k+1}} \;=\; \mathbb{E}_{t_k}\big[\bar{Y}^\pi_{t_{k+1}}\big] + \int_{t_k}^{t_{k+1}} \tilde{\mathcal{Z}}^\pi_t dW_t + \int_{t_k}^{t_{k+1}}\int_A \tilde{\mathcal{U}}^\pi_t(a)\tilde{\mu}(dt,da), \quad k<n,$$

and we can then define the continuous-time processes $\bar{Y}^\pi$ and $\bar{\mathcal{Y}}^\pi$ by:

$$\begin{aligned}
\bar{\mathcal{Y}}^\pi_t \;&=\; \bar{Y}^\pi_{t_{k+1}} + (t_{k+1}-t)f(\bar{X}^\pi_{t_k}, I_{t_k}, \bar{\mathcal{Y}}^\pi_{t_k}) \qquad (5.3)\\
&\quad - \int_t^{t_{k+1}} \tilde{\mathcal{Z}}^\pi_t dW_t - \int_t^{t_{k+1}}\int_A \tilde{\mathcal{U}}^\pi_t(a)\tilde{\mu}(dt,da), \qquad t\in[t_k,t_{k+1}),\\
\bar{Y}^\pi_t \;&=\; \bar{Y}^\pi_{t_{k+1}} + (t_{k+1}-t)f(\bar{X}^\pi_{t_k}, I_{t_k}, \bar{\mathcal{Y}}^\pi_{t_k}) \qquad (5.4)\\
&\quad - \int_t^{t_{k+1}} \tilde{\mathcal{Z}}^\pi_t dW_t - \int_t^{t_{k+1}}\int_A \tilde{\mathcal{U}}^\pi_t(a)\tilde{\mu}(dt,da), \qquad t\in(t_k,t_{k+1}],
\end{aligned}$$

for $k=0,\ldots,n-1$. Denote by $\delta Y^\pi_t = Y^\pi_t - \bar{Y}^\pi_t$, $\delta\mathcal{Y}^\pi_t = \mathcal{Y}^\pi_t - \bar{\mathcal{Y}}^\pi_t$, $\delta\mathcal{Z}^\pi_t = \mathcal{Z}^\pi_t - \tilde{\mathcal{Z}}^\pi_t$, $\delta\mathcal{U}^\pi_t = \mathcal{U}^\pi_t - \tilde{\mathcal{U}}^\pi_t$ and $\delta f_t = f(X_t, I_t, \mathcal{Y}^\pi_t) - f(\bar{X}^\pi_{t_k}, I_{t_k}, \bar{\mathcal{Y}}^\pi_{t_k})$ for $t\in[t_k,t_{k+1})$. Recalling (3.6) and (5.3), we have by Itô's formula:

$$\begin{aligned}
\Delta_t \;&:=\; \mathbb{E}_{t_k}\Big[|\delta\mathcal{Y}^\pi_t|^2 + \int_t^{t_{k+1}} |\delta\mathcal{Z}^\pi_s|^2 ds + \int_t^{t_{k+1}}\int_A |\delta\mathcal{U}^\pi_s(a)|^2\lambda(da)ds\Big]\\
&=\; \mathbb{E}_{t_k}\big[|\delta Y^\pi_{t_{k+1}}|^2\big] + \mathbb{E}_{t_k}\Big[\int_t^{t_{k+1}} 2\delta\mathcal{Y}^\pi_s\delta f_s\Big] ds
\end{aligned}$$

for all $t\in[t_k,t_{k+1})$. By the Lipschitz continuity of f in **(H2)** and Young inequality, we then have:

$$\begin{aligned}
\Delta_t \;&\leq\; \mathbb{E}_{t_k}\big[|\delta Y^\pi_{t_{k+1}}|^2\big] + \mathbb{E}_{t_k}\Big[\int_t^{t_{k+1}} \eta|\delta\mathcal{Y}^\pi_s|^2 ds + \frac{C}{\eta}\pi|\delta\mathcal{Y}^\pi_{t_k}|^2\Big]\\
&\quad + \frac{C}{\eta}\mathbb{E}_{t_k}\Big[\int_t^{t_{k+1}} \big(|X_s - \bar{X}^\pi_{t_k}|^2 + |I_s - I_{t_k}|^2 + |\mathcal{Y}^\pi_s - \mathcal{Y}^\pi_{t_k}|^2\big) ds\Big].
\end{aligned}$$

From Gronwall's lemma, and by taking η large enough, this yields for all $k\leq n-1$:

$$\mathbb{E}_{t_k}\Big[|\delta\mathcal{Y}^\pi_{t_k}|^2\Big] \;\leq\; e^{C|\pi|}\mathbb{E}_{t_k}\big[|\delta Y^\pi_{t_{k+1}}|^2\big] + CB_k \qquad (5.5)$$

where

$$\begin{aligned}
B_k \;&=\; \mathbb{E}_{t_k}\Big[\int_{t_k}^{t_{k+1}} \big(|X_s - \bar{X}^\pi_{t_k}|^2 + |I_s - I_{t_k}|^2 + |\mathcal{Y}^\pi_s - \mathcal{Y}^\pi_{t_k}|^2\big) ds\Big]\\
&\leq\; C|\pi|\Big(\mathbb{E}_{t_k}\big[\sup_{s\in[t_k,t_{k+1}]} |X_s - \bar{X}^\pi_{t_k}|^2\big] + |\pi|(1+|X_{t_k}|)\Big), \qquad (5.6)
\end{aligned}$$

by (2.5) and Proposition 3.3. Now, by definition of $Y^\pi_{t_{k+1}}$ and $\bar{Y}^\pi_{t_{k+1}}$, we have

$$|\delta Y^\pi_{t_{k+1}}|^2 \;\leq\; \operatorname*{ess\,sup}_{a\in A} \mathbb{E}_{t_{k+1},a}\big[|\delta\mathcal{Y}^\pi_{t_{k+1}}|^2\big]. \qquad (5.7)$$

By plugging (5.6), (5.7) into (5.5), taking conditional expectation with respect to $I_{t_k} = a$, and taking essential supremum over a, we obtain:

$$
\begin{aligned}
\operatorname*{ess\,sup}_{a\in A} \mathbb{E}_{t_k,a}\Big[|\delta\mathcal{Y}^\pi_{t_k}|^2\Big] \;\; &\leq \;\; e^{C|\pi|}\operatorname*{ess\,sup}_{a\in A} \mathbb{E}_{t_k,a}\big[\operatorname*{ess\,sup}_{a\in A} \mathbb{E}_{t_{k+1},a}\big[|\delta\mathcal{Y}^\pi_{t_{k+1}}|^2\big] \\
&\qquad + C|\pi|\Big(\operatorname*{ess\,sup}_{a\in A} \mathbb{E}_{t_k,a}\big[\sup_{s\in[t_k,t_{k+1}]} |X_s - \bar{X}^\pi_{t_k}|^2\big] + |\pi|(1+|X_{t_k}|)\Big).
\end{aligned}
$$

By taking conditional expectation with respect to $\mathcal{F}_{t_{k-1}}$, and $I_{t_{k-1}} = a$, taking essential supremum over a in the above inequality, and iterating this backward procedure until time $t_0 = 0$, we obtain:

$$
\begin{aligned}
\mathcal{E}^\pi_k(\mathcal{Y}) \;\; &\leq \;\; e^{C|\pi|}\mathcal{E}^\pi_{k+1}(\mathcal{Y}) + C|\pi|\big(\mathcal{E}^\pi_k(X) + |\pi|(1+\mathbb{E}[|X_{t_k}|])\big) \\
&\leq \;\; e^{C|\pi|}\mathcal{E}^\pi_{k+1}(\mathcal{Y}) + C|\pi|^2, \qquad k \leq n-1,
\end{aligned}
\tag{5.8}
$$

where we recall the auxiliary error control $\mathcal{E}^\pi_k(X)$ on X in (2.2) and its estimate in Lemma 2.1, and set:

$$
\mathcal{E}^\pi_k(\mathcal{Y}) \;\; := \;\; \mathbb{E}\Big[\operatorname*{ess\,sup}_{a\in A} \mathbb{E}_{t_1,a}\big[\ldots \operatorname*{ess\,sup}_{a\in A} \mathbb{E}_{t_k,a}\big[|\delta\mathcal{Y}^\pi_{t_k}|^2\big]\ldots\big]\Big].
$$

By a direct induction on (5.8), and recalling that $n|\pi|$ is bounded, we get

$$
\begin{aligned}
\mathcal{E}^\pi_k(\mathcal{Y}) \;\; &\leq \;\; C\big(\mathcal{E}^\pi_n(\mathcal{Y}) + |\pi|\big) \\
&\leq \;\; C\big(\mathcal{E}^\pi_n(X) + |\pi|\big) \; \leq \; C|\pi|,
\end{aligned}
$$

since g is Lipschitz, and using again the estimate in Lemma 2.1. Observing that $\mathbb{E}[|\delta Y^\pi_{t_k}|^2]$, $\mathbb{E}[|\delta\mathcal{Y}^\pi_{t_k}|^2] \leq \mathcal{E}^\pi_k(\mathcal{Y})$, we get the estimate:

$$
\max_{k\leq n} \mathbb{E}\big[|Y^\pi_{t_k} - \bar{Y}^\pi_{t_k}|^2\big] + \mathbb{E}\big[|\mathcal{Y}^\pi_{t_k} - \bar{\mathcal{Y}}^\pi_{t_k}|^2\big] \;\; \leq \;\; C|\pi|.
$$

Moreover, by Proposition 3.3, we have

$$
\begin{aligned}
\sup_{t\in[t_k,t_{k+1})} \mathbb{E}\Big[|\mathcal{Y}^\pi_t - \mathcal{Y}^\pi_{t_k}|^2\Big] + \sup_{t\in(t_k,t_{k+1}]} \mathbb{E}\Big[|Y^\pi_t - Y^\pi_{t_{k+1}}|^2\Big] \;\; &\leq \;\; C(1+\mathbb{E}[|X_{t_k}|])|\pi| \\
&\leq \;\; C(1+|X_0|)|\pi|.
\end{aligned}
$$

This implies finally that:

$$
\begin{aligned}
\sup_{s\in(t_k,t_{k+1}]} \mathbb{E}\Big[|Y^\pi_t - \bar{Y}^\pi_{t_{k+1}}|^2\Big] \;\; &\leq \;\; 2\sup_{s\in(t_k,t_{k+1}]} \mathbb{E}\Big[|Y^\pi_t - Y^\pi_{t_{k+1}}|^2\Big] + 2\mathbb{E}\Big[|Y^\pi_{t_{k+1}} - \bar{Y}^\pi_{t_{k+1}}|^2\Big] \\
&\leq \;\; C|\pi|,
\end{aligned}
$$

as well as

$$
\begin{aligned}
\sup_{s\in[t_k,t_{k+1})} \mathbb{E}\Big[|Y^\pi_t - \bar{\mathcal{Y}}^\pi_{t_k}|^2\Big] \;\; &\leq \;\; 2\sup_{s\in[t_k,t_{k+1})} \mathbb{E}\Big[|Y^\pi_t - \mathcal{Y}^\pi_{t_k}|^2\Big] + 2\mathbb{E}\Big[|\mathcal{Y}^\pi_{t_k} - \bar{\mathcal{Y}}^\pi_{t_k}|^2\Big] \\
&\leq \;\; C|\pi|.
\end{aligned}
$$

□

5.2 Approximate optimal control

We now consider the special case where $f(x, a)$ does not depend on y, so that the discrete time scheme (1.4) is an approximation for the value function of the stochastic control problem:

$$\begin{aligned} V_0 &:= \sup_{\alpha \in \mathcal{A}} J(\alpha) \;=\; Y_0, \qquad (5.9) \\ J(\alpha) &= \mathbb{E}\Big[\int_0^T f(X_t^\alpha, \alpha_t)dt + g(X_T^\alpha) \Big], \end{aligned}$$

where $\mathcal{A}$ is the set of $\mathbb{G}$-adapted control processes α valued in A, and X^α is the controlled diffusion in $\mathbb{R}^d$:

$$X_t^\alpha \;=\; X_0 + \int_0^t b(X_s^\alpha, \alpha_s)ds + \int_0^t \sigma(X_s^\alpha, \alpha_s)dW_s, \quad 0 \leq t \leq T.$$

(Here $\mathbb{G} = (\mathcal{G}_t)_{0 \leq t \leq T}$ denotes some filtration under which W is a standard Brownian motion). Let us now define the discrete time version of (5.9). We introduce the set $\mathcal{A}^\pi$ of discrete time processes $\alpha = (\alpha_{t_k})_k$ with α_{t_k} $\mathcal{G}_{t_k}$-measurable, and valued in A. For each $\alpha \in \mathcal{A}^\pi$, we consider the controlled discrete time process $(X_{t_k}^{\pi,\alpha})_k$ of Euler type defined by:

$$X_{t_k}^{\pi,\alpha} \;=\; X_0 + \sum_{j=0}^{k-1} b(X_{t_j}^{\pi,\alpha}, \alpha_{t_j})\Delta t_j + \sum_{j=0}^{k-1} \sigma(X_{t_j}^{\pi,\alpha}, \alpha_{t_j})\Delta W_{t_j}, \quad k \leq n,$$

where $\Delta W_{t_j} = W_{t_{j+1}} - W_{t_j}$, and the gain functional:

$$J^\pi(\alpha) \;=\; \mathbb{E}\Big[\sum_{k=0}^{n-1} f(X_{t_k}^{\pi,\alpha}, \alpha_{t_k})\Delta t_k + g(X_{t_n}^{\pi,\alpha}) \Big].$$

Given any $\alpha \in \mathcal{A}^\pi$, we define its continuous time piecewise-constant interpolation $\alpha \in \mathcal{A}$ by setting: $\alpha_t = \alpha_{t_k}$, for $t \in [t_k, t_{k+1})$ (by misuse of notation, we keep the same notation α for the discrete time and continuous time interpolation). By standard arguments similar to those for Euler scheme of SDE, there exists some positive constant C such that for all $\alpha \in \mathcal{A}^\pi$, $k \leq n-1$:

$$\mathbb{E}\Big[\sup_{t \in [t_k, t_{k+1}]} \big| X_t^\alpha - X_{t_k}^{\pi,\alpha} \big|^2 \Big] \;\leq\; C|\pi|,$$

from which we easily deduce by Lipschitz property of f and g:

$$\big| J(\alpha) - J^\pi(\alpha) \big| \;\leq\; C|\pi|^{\frac{1}{2}}, \quad \forall \alpha \in \mathcal{A}^\pi. \qquad (5.10)$$

Let us now consider at each time step $k \leq n-1$, the function $\hat{a}_k(x)$ which attains the supremum over $a \in A$ of $\bar{\vartheta}_k^\pi(x, a)$ in the scheme (5.2), so that:

$$\bar{v}_k^\pi(x) \;=\; \bar{\vartheta}_k^\pi\big(x, \hat{a}_k(x)\big), \quad k = 0, \ldots, n-1.$$

Let us define the process $(\hat{X}_{t_k}^\pi)_k$ by: $\hat{X}_0^\pi = X_0$,

$$\hat{X}_{t_{k+1}}^\pi \;=\; \hat{X}_{t_k}^\pi + b(\hat{X}_{t_k}^\pi, \hat{a}_k(\hat{X}_{t_k}^\pi))\Delta t_k + \sigma(\hat{X}_{t_k}^\pi, \hat{a}_k(\hat{X}_{t_k}^\pi))\Delta W_{t_k}, \quad k \leq n-1,$$

and notice that $\hat{X}^\pi = X^{\pi,\hat{\alpha}}$, where $\hat{\alpha} \in \mathcal{A}^\pi$ is a feedback control defined by:

$$\hat{\alpha}_{t_k} \;=\; \hat{a}_k(\hat{X}^\pi_{t_k}) \;=\; \hat{a}_k(X^{\pi,\hat{\alpha}}_{t_k}), \quad k = 0, \ldots, n.$$

Next, we observe that the conditional law of $\bar{X}^\pi_{t_{k+1}}$ given $(\bar{X}^\pi_{t_k} = x, I_{t_k} = \hat{a}_k(\bar{X}^\pi_{t_k}) = \hat{a}_k(x))$ is the same than the conditional law of $X^{\pi,\hat{\alpha}}_{t_{k+1}}$ given $X^{\pi,\hat{\alpha}}_{t_k} = x$, for $k \leq n-1$, and thus the induction step in the scheme (5.1) or (5.2) reads as:

$$\bar{v}^\pi_k(X^{\pi,\hat{\alpha}}_{t_k}) \;=\; \mathbb{E}\Big[\bar{v}^\pi_{k+1}(X^{\pi,\hat{\alpha}}_{t_{k+1}})\big|X^{\pi,\hat{\alpha}}_{t_k}\Big] + f(X^{\pi,\hat{\alpha}}_{t_k}, \hat{\alpha}_{t_k})\Delta t_k, \quad k \leq n-1.$$

By induction, and law of iterated conditional expectations, we then get:

$$\bar{Y}^\pi_0 \;=\; \bar{v}^\pi_0(X_0) \;=\; J^\pi(\hat{\alpha}). \tag{5.11}$$

Consider the continuous time piecewise-constant interpolation $\hat{\alpha} \in \mathcal{A}$ defined by: $\hat{\alpha}_t = \hat{\alpha}_{t_k}$, for $t \in [t_k, t_{k+1})$. By (5.9), (5.10), (5.11), and Corollary 5.1, we finally obtain:

$$\begin{aligned} 0 \;\leq\; V_0 - J(\hat{\alpha}) \;&=\; Y_0 - \bar{Y}^\pi_0 + J^\pi(\hat{\alpha}) - J(\hat{\alpha}) \\ &\leq\; C|\pi|^{\frac{1}{6}} + C|\pi|^{\frac{1}{2}} \;\leq\; C|\pi|^{\frac{1}{6}}, \end{aligned}$$

for $|\pi| \leq 1$. In other words, for any small $\varepsilon > 0$, we obtain an ε-approximate optimal control $\hat{\alpha}$ for the stochastic control problem (5.9) by taking $|\pi|$ of order ε^6.

References

[1] Bally V. and G. Pagès (2003): "Error analysis of the quantization algorithm for obstacle problems", *Stochastic processes and their applications*, 106, 1-40.

[2] Barles G. (1994): *Solutions de viscosité des équations d'Hamilton-Jacobi*, Mathématiques et Applications, Springer Verlag.

[3] Barles G., Buckdahn R. and E. Pardoux (1997): "BSDEs and integral-partial differential equations", *Stochastics and Stochastic Reports*, 60, 57-83

[4] Barles G. and E.R. Jacobsen (2007): "Error bounds for monotone approximation schemes for parabolic Hamilton-Jacobi-Bellman equations", *Math. Computation*, 76, 1861-1893.

[5] Bonnans F. and H. Zidani (2003): "Consistency of generalized finite difference schemes for the stochastic HJB equation", *SIAM J. Numer. Anal.*, 41, 1008-1021.

[6] Bouchard B. and J.F. Chassagneux (2008): "Discrete time approximation for continuously and discretely reflected BSDEs", *Stochastic Processes and Applications*, 118, 2269-2293.

[7] Bouchard B. and R. Elie (2008): "Discrete time approximation of decoupled FBSDE with jumps", *Stochastic Processes and Applications*, 118 (1), 53-75.

[8] Bouchard B. and N. Touzi (2004): "Discrete time approximation and Monte-Carlo simulation of backward stochastic differential equations", *Stochastic Processes and their Applications*, 111 (2), 175-206.

[9] Crandall M., Ishii H. and P.L. Lions (1992) : "User's guide to viscosity solutions of second order partial differential equations", *Bull. Amer. Math. Soc.*, 27, 1-67.

[10] Fahim A., Touzi N. and X. Warin (2011): "A Probabilistic Numerical Scheme for Fully Nonlinear PDEs", *The Annals of Applied Probability*, 21 (4), 1322-1364.

[11] Fleming W. H. and H. M. Soner (2006), *Controlled Markov Processes and Viscosity Solutions*, 2nd edition, Springer-Verlag.

[12] Gobet E., Lemor J.P. and X. Warin (2006): "Rate of convergence of empirical regression method for solving generalized BSDE", *Bernoulli*, 12 (5), 889-916.

[13] Kharroubi I. and H. Pham (2012): "Feynman-Kac representation for Hamilton-Jacobi-Bellman IPDE", *Preprint*.

[14] Kharroubi I., Langrené N. and H. Pham (2013): "Numerical algorithm for fully nonlinear HJB equations: an approach by control randomization".

[15] Kloeden P. and E. Platen (1992): *Numerical solution of stochastic differential equations*, Springer.

[16] Krylov N.V. (1999): "Approximating value functions for controlled degenerate diffusion processes by using piece-wise constant policies", *Electronic Journal of Probability*, 4, 1-19.

[17] Krylov N. V. (2000): "On the rate of convergence of finite difference approximations for Bellman's equations with variable coefficients", *Probability Theory and Related Fields*, 117, 1-16.

[18] Kushner H. and P. Dupuis (1992): Numerical methods for stochastic control problems in continuous time, Springer Verlag.

[19] Pham H. (2009): *Continuous-time stochastic control and optimization with financial applications*, Springer, Series SMAP, Vol, 61.

[20] Tan X. (2013): "A splitting method for fully nonlinear degenerate parabolic PDEs", *Electron. J. Probab.*, 18, n. 15, 1-24.

[21] Zhang J. (2004): "A numerical scheme for BSDEs", *Annals of Applied Probability*, 14 (1), 459-488.